



(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

**EP 0 419 182 B1**

(12)

**EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention  
of the grant of the patent:  
19.01.2000 Bulletin 2000/03

(51) Int. Cl.<sup>7</sup>: **C12N 15/51, C07K 4/02**  
// A61K39/29, G01N33/576,  
C12Q1/68, C12Q1/70

(21) Application number: 90310149.1

(22) Date of filing: 17.09.1990

(54) **New HCV Isolate J-1**

HCV-Isolat J-1

Isolate d'HCV J-1

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB GR IT LI LU NL SE

(30) Priority: 15.09.1989 US 408045  
21.12.1989 US 456142

(43) Date of publication of application:  
27.03.1991 Bulletin 1991/13

(60) Divisional application:  
99101746.8 / 0 939 128

(73) Proprietors:  
• Oya, Akira, Dr.,  
Director General of the National Institute of  
Health of Japan,  
on behalf of The National Institute of Japan  
Shinagawa-ku, Tokyo 141 (JP)  
• CHIRON CORPORATION  
Emeryville, California 94608 (US)

(72) Inventors:  
• Miyamura, Tatsuo  
Tokyo 168 (JP)

- Salto, Izumu  
Tokyo 151 (JP)
- Houghton, Michael  
Danville, California 94526 (US)
- Weiner, Amy J.  
Benicia, California 94510 (US)
- Han, Jang  
Lafayette, California 94549 (US)
- Kolberg, Janice A.  
Richmond, California 94804 (US)
- Cha, Tai-An  
San Ramon, California 94583 (US)
- Irvine, Bruce Duncan  
Concord, California 94519 (US)

(74) Representative:  
Goldin, Douglas Michael et al  
J.A. KEMP & CO.  
14 South Square  
Gray's Inn  
London WC1R 5LX (GB)

(56) References cited:  
EP-A- 0 318 216                      EP-A- 0 388 232

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

**EP 0 419 182 B1**

**Description****Technical Field**

5 [0001] The present invention relates to new isolates of the viral class Hepatitis C, polypeptides, polynucleotides and antibodies derived therefrom, as well as the use of such polypeptides, polynucleotides and antibodies in assays (e.g., immunoassays, nucleic acid hybridization assays, etc.) and in the production of viral polypeptides.

**Background**

10 [0002] Non-A, Non-B hepatitis (NANBH) is a transmissible disease or family of diseases that are believed to be viral-induced, and that are distinguishable from other forms of viral-associated liver diseases, including that caused by the known hepatitis viruses, i.e., hepatitis A virus (HAV), hepatitis B virus (HBV), and delta hepatitis virus (HDV), as well as the hepatitis induced by cytomegalovirus (CMV) or Epstein-Barr virus (EBV). NANBH was first identified in transfused  
15 individuals. Transmission from man to chimpanzee and serial passage in chimpanzees provided evidence that NANBH is due to a transmissible infectious agent or agents. Epidemiologic evidence is suggestive that there may be three types of NANBH: the water-borne epidemic type; the blood or needle associated type; and the sporadically occurring (community acquired) type. However, until recently, no transmissible agent responsible for NANBH had not been identified.

[0003] Clinical diagnosis and identification of NANBH has been accomplished primarily by exclusion of other viral  
20 markers. Among the methods used to detect putative NANBH antigens and antibodies are agar-gel diffusion, counter-immunoelectrophoresis, immunofluorescence microscopy, immune electron microscopy, radioimmunoassay, and enzyme-linked immunosorbent assay. However, none of these assays has proved to be sufficiently sensitive, specific, and reproducible to be used as a diagnostic test for NANBH.

[0004] Until recently there has been neither clarity nor agreement as to the identity or specificity of the antigen antibody systems associated with agents of NANBH. It is possible that NANBH is caused by more than one infectious agent  
25 and unclear what the serological assays detect in the serum of patients with NANBH.

[0005] In the past, a number of candidate NANBH agents were postulated. See, e.g., Prince (1983) Ann. Rev. Microbiol. 37:217; Feinstone & Hoofnagle (1984) New Eng. J. Med. 311:185; Overby (1985) Curr. Hepatol. 5:49; Overby (1986) Curr. Hepatol. 6:65; Overby (1987) Curr. Hepatol. 7:35; and Iwarson (1987) British Med. J. 295:946. However, there  
30 is no proof that any of these candidates represent the etiological agent of NANBH.

[0006] In 1987, Houghton et al. cloned the first virus definitively linked to NANBH. See, e.g., EPO Pub. No. 318,216; Houghton et al., Science 244:359 (1989). Houghton et al. described therein the cloning of an isolate from a new viral class, hepatitis C virus (HCV), the prototype isolate described therein being named "HCV1". HCV is a Flavi-like virus,  
35 with an RNA genome. Houghton et al. described the production of recombinant proteins from HCV sequences that are useful as diagnostic reagents, as well as polynucleotides useful in diagnostic hybridization assays and in the cloning of additional HCV isolates.

[0007] The demand for sensitive, specific methods for screening and identifying carriers of NANBH and NANBH contaminated blood or blood products is significant. Post-transfusion hepatitis (PTH) occurs in approximately 10% of transfused patients, and NANBH accounts for up to 90% of these cases. There is a frequent progression to chronic liver  
40 damage (25-55%).

[0008] Patient care as well as the prevention of transmission of NANBH by blood and blood products or by close personal contact require reliable diagnostic and prognostic tools to detect nucleic acids, antigens and antibodies related to NANBH. In addition, there is also a need for effective vaccines and immunotherapeutic therapeutic agents for the prevention and/or treatment of the disease.

45 [0009] While at least one HCV isolate has been identified which is useful in meeting the above needs, additional isolates, particularly those with divergent a genome, may prove to have unique applications.

**Summary of the invention**

50 [0010] New isolates of HCV has been characterized from Japanese blood donors who have been implicated as NANBH carriers. These isolates exhibit nucleotide and amino acid sequence heterogeneity with respect to the prototype isolate, HCV1, in several viral domains. It is believed that these distinct sequences are of importance, particularly in diagnostic assays and in vaccine development.

[0011] In one embodiment, the present invention provides a polynucleotide in substantially isolated form comprising  
55 a nucleotide sequence of at least 8 nucleotides from a J-1 HCV isolate, said J-1 HCV isolate having at least 90% nucleotide sequence homology with the J-1 sequence of any one of Figures 7 to 10 or 13 to 18, wherein said nucleotide sequence of at least 8 nucleotides is distinct from the nucleotide sequence of HCV isolate HCV-1.

[0012] In another embodiment of the present invention provides a purified polypeptide comprising a amino acid

sequence which:

- (a) is encoded by a nucleotide sequence of the invention, said coding being in frame with the corresponding amino acid sequences set out in Figures 7 to 10 or 13 to 18;
- (b) comprises an antigenic determinant; and
- (c) contains at least one amino acid substitution compared to the amino acids encoded by the corresponding portion of the HCV-1 sequence.

[0013] Yet another embodiment of the present invention provides a polypeptide of the invention immobilized on the solid support.

[0014] In a further embodiment of the present invention, an immunoassay for detecting the presence of anti-HCV antibodies in a test sample is provided comprising: (a) incubating the test sample under conditions that allow the formation of antigen-antibody complexes with polypeptide of the invention, wherein the polypeptide is not immunologically cross-reactive with HCV-1 and (b) detecting any antigen-antibody complexes formed.

[0015] Yet another embodiment of the present invention provides a method of detecting HCV polynucleotides in a test sample comprising: (a) providing a polypeptide of the invention as a probe; (b) contacting the test sample and the probe under conditions that allow for the formation of a polynucleotide duplex between the probe and its complement in the absence of substantial polynucleotide duplex formation between the probe and non-HCV polynucleotide sequences present in the test sample; and (c) detecting any polynucleotide duplexes comprising the probe.

[0016] These and other embodiments of the present invention will be readily apparent to this of ordinary skill in the art in view of the following description.

#### Brief description of the Figures

[0017]

Figure 1 shows the consensus sequence of the coding strand of a fragment from the J7 C/E domain with the heterogeneities.

Figure 2 shows the consensus sequence of the coding strand of a fragment from the J1 E domain with the heterogeneities.

Figure 3 shows the consensus sequence of the coding strand of a fragment of the J1 E/NS1 domain with the heterogeneities.

Figure 4 shows the consensus sequence of the coding strand of a fragment from the J1 NS3 domain with the heterogeneities.

Figure 5 shows the consensus sequence of the coding strand of a fragment from the J1 NS5 domain with the heterogeneities.

Figure 6 shows the homology of the J7 C/E consensus sequence with the nucleotide sequence of the same domain from HCV1.

Figure 7 shows the homology of the J1 E consensus sequence with the nucleotide sequence of the same domain from HCV1.

Figure 8 shows the homology of the J1 E/NS1 consensus sequence with the nucleotide sequence of the same domain from HCV1.

Figure 9 shows the homology of the J1 NS3 consensus sequence with the nucleotide sequence of the same domain from HCV1.

Figure 10 shows the homology of the J1 NS5 consensus sequence with the nucleotide sequence of the same domain from HCV1.

Figure 11 shows the putative genomic organization of the HCV1 genome.

Figure 12 shows the nucleotide sequence of the ORF of HCV1. In the figure nucleotide number 1 is the first A of the putative initiating methionine of the large ORF; nucleotides upstream of this nucleotide are numbered with negative numbers.

Figure 13 shows the consensus sequence of the coding strand of a fragment from the J1 NS1 domain (J1 1519) with the nucleotide sequence of the same domain from HCV1. Also shown are the amino acids encoded therein.

Figure 14 shows a composite of the consensus sequence from the core to the NS1 domain of J1 with the nucleotide sequence of the same domain from HCV1. Also shown are the amino acids encoded therein.

Figure 15 shows a consensus sequence of the coding strand of the NS1 domain of J1, as determined in Example IV. Also shown are the nucleotide sequence of the same domain from HCV1, and the amino acids encoded in the HCV1 and J1 sequences.

Figure 16 shows a consensus sequence of a coding strand of the C200 region of the NS3-NS4 domain of J1. Also

shown are the nucleotide sequence of the same domain from HCV1. Also shown are the amino acids encoded in the sequences.

Figure 17 shows a consensus sequence of the coding strand of the NS1 domain of J1, as determined in Example V. Also shown are the nucleotide sequence of the same domain from HCV1, and the amino acids encoded in the sequences.

Figure 18 shows a consensus sequence of the coding strand of the untranslated and core domains of J1. Also shown are the nucleotide sequence of the same domain from HCV1, and the amino acids encoded in the sequences.

#### 10 Detailed Description of the Invention

[0018] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA techniques, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Maniatis, Fritsch & Sambrook, MOLECULAR CLONING; A LABORATORY MANUAL (1982); DNA CLONING, VOLUMES I AND II (D.N. Glover ed. 1985); OLIGONUCLEOTIDE SYNTHESIS (M.J. Gait ed. 1984); NUCLEIC ACID HYBRIDIZATION (B.D. Names & S.J. Higgins eds. 1984); TRANSCRIPTION AND TRANSLATION (B.D. Hames & S.J. Higgins eds. 1984); ANIMAL CELL CULTURE (R.I. Freshney ed. 1986); IMMOBILIZED CELLS AND ENZYMES (IRL Press, 1986); B. Perbal, A PRACTICAL GUIDE TO MOLECULAR CLONING (1984); the series, METHODS IN ENZYMOLOGY (Academic Press, Inc.); GENE TRANSFER VECTORS FOR MAMMALIAN CELLS (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory), Methods in Enzymology Vol. 154 and Vol. 155 (Wu and Grossman, and Wu, eds., respectively), Mayer and Walker, eds. (1987), IMMUNOCHEMICAL METHODS IN CELL AND MOLECULAR BIOLOGY (Academic Press, London), Scopes, (1987), PROTEIN PURIFICATION: PRINCIPLES AND PRACTICE, Second Edition (Springer-Verlag, N.Y.), and HANDBOOK OF EXPERIMENTAL IMMUNOLOGY, VOLUMES I-IV (D.M. Weir and C. C. Blackwell eds 1986). All patents, patent applications, and other publications mentioned herein, both supra and infra, are hereby incorporated herein by reference.

[0019] The term "hepatitis C virus" has been reserved by workers in the field for an heretofore unknown etiologic agent of NANBH. Accordingly, as used herein, "hepatitis C virus" (HCV) refers to an agent causative of NANBH, which was formerly referred to as NANBV and/or BB-NANBV from the class of the prototype isolate, HCV1, described by Houghton et al. See, e.g., EPO Pub. No. 318,216 and U.S. Patent App. Serial No. 355,002, filed 19 May 1989 (available in non-U.S. applications claiming priority therefrom), the disclosures of which are incorporated herein by reference. The nucleotide sequence and putative amino acid sequence of HCV1 is shown in Figure 6. The terms HCV, NANBV, and BB-NANBV are used interchangeably herein. As an extension of this terminology, the disease caused by HCV, formerly called NANB hepatitis (NANBH), is called hepatitis C. The terms NANBH and hepatitis C may be used interchangeably herein. The term "HCV", as used herein, denotes a viral species of which pathogenic strains cause NANBH, as well as attenuated strains or defective interfering particles derived therefrom.

[0020] HCV is a Flavi-like virus. The morphology and composition of Flavivirus particles are known, and are discussed by Brinton (1986) THE VIRUSES: THE TOGAVIRIDAE AND FLAVIVIRIDAE (Series eds. Fraenkel-Conrat and Wagner, vol eds. Schlesinger and Schlesinger, Plenum Press), p.327-374. Generally, with respect to morphology, Flaviviruses contain a central nucleocapsid surrounded by a lipid bilayer. Virions are spherical and have a diameter of about 40-50 nm. Their cores are about 25-30 nm in diameter. Along the outer surface of the virion envelope are projections that are about 5-10 nm long with terminal knobs about 2 nm in diameter.

[0021] The HCV genome is comprised of RNA. It is known that RNA containing viruses have relatively high rates of spontaneous mutation, i.e., reportedly on the order of  $10^{-3}$  to  $10^{-4}$  per incorporated nucleotide. Therefore, there are multiple strains, which may be virulent or avirulent, within the HCV class or species.

[0022] It is believed that the genome of HCV isolates is comprised of a single ORF of approximately 9,000 nucleotides to approximately 12,000 nucleotides, encoding a polyprotein similar in size to that of HCV1, an encoded polyprotein of similar hydrophobic and antigenic character to that of HCV1, and the presence of co-linear peptide sequences that are conserved with HCV1. In addition, the genome is believed to be a positive-stranded RNA.

[0023] Isolates of HCV comprise epitopes that are immunologically cross-reactive with epitopes in the HCV1 genome. At least some of these are epitopes unique to HCV when compared to other known Flaviviruses. The uniqueness of the epitope may be determined by its immunological reactivity with anti-HCV antibodies and lack of immunological reactivity with antibodies to other Flavivirus species. Methods for determining immunological reactivity are known in the art, for example, by radioimmunoassay, by ELISA assay, by hemagglutination, and several examples of suitable techniques for assays are provided herein.

[0024] It is also expected that the overall homology of HCV isolates and HCV1 genomes at the nucleotide level probably will be about 40% or greater, probably about 60% or greater, and even more probably about 80% to about 90% or greater. In addition that there are many corresponding contiguous sequences of at least about 13 nucleotides that are

fully homologous. The correspondence between the sequence from a new isolate and the HCV1 sequence can be determined by techniques known in the art. For example, they can be determined by a direct comparison of the sequence information of the polynucleotide from the new isolate and HCV1 sequences. Alternatively, homology can be determined by hybridization of the polynucleotides under conditions which form stable duplexes between homologous regions (for example, those which would be used prior to S<sub>1</sub> digestion), followed by digestion with single-stranded specific nuclease(s), followed by size determination of the digested fragments.

[0025] Because of the evolutionary relationship of the strains or isolates of HCV, putative HCV strains or isolates are identifiable by their homology at the polypeptide level. Thus, new HCV isolates are expected to be more than about 40% homologous, probably more than about 70% homologous, and even more probably more than about 80% homologous, and possibly even more than about 90% homologous at the polypeptide level. The techniques for determining amino acid sequence homology are known in the art. For example, the amino acid sequence may be determined directly and compared to the sequences provided herein. Alternatively the nucleotide sequence of the genomic material of the putative HCV may be determined, the amino acid sequence encoded therein can be determined, and the corresponding regions compared.

[0026] The ORF of HCV1 is shown in Figure 12. The non-structural, core, and envelope domains of the polyprotein have been predicted for HCV1 (Figure 5). The "C", or core, polypeptide is believed to be encoded from the 5' terminus to about nucleotide 345 of HCV1. The putative "E", or envelope, domain of HCV1 is believed to be encoded from about nucleotide 346 to about nucleotide 1050. Putative NS1, or non-structural one domain, is thought to be encoded from about nucleotide 1051 to about nucleotide 1953. For the remaining domains, putative NS2 is thought to be encoded from about nucleotide 1954 to about nucleotide 3018, putative NS3 from about nucleotide 3019 to about nucleotide 4950, putative NS4 from about nucleotide 4951 to about nucleotide 6297, and putative NS5 from about nucleotide 6298 to the 3' terminus respectively. The above boundaries are approximations based on an analysis of the ORF. The exact boundaries can be determined by those skilled in the art in view of the disclosure herein.

[0027] "HCV/J1" or "J1" and "HCV/J7" or "J7" refer to new HCV isolates characterized by the nucleotide sequence disclosed herein, as well as related isolates that are substantially homologous thereto; i.e., at least about 90% or about 95% at the nucleotide level. It is believed that the sequences disclosed herein characterize an HCV subclass that is predominant in Japan and other Asian and/or Pacific rim countries. Additional J1 and J7 isolates can be obtained in view of the disclosure herein and EPO Pub. No. 318,216. In particular, the J1 and J7 nucleotide sequences disclosed herein, as well as the HCV1 sequences in Figure 12, can be used as primers or probes to clone additional domains of J1, J7, or additional isolates.

[0028] As used herein, a nucleotide sequence "from" a designated sequence or source refers to a nucleotide sequence that is homologous (i.e., identical) to or complementary to the designated sequence or source, or a portion thereof. The J1 sequences provided herein are a minimum of 15 nucleotides preferably 20 nucleotides or longer. The maximum length is the complete viral genome.

[0029] In some aspects of the invention, the sequence of the region from which the polynucleotide is derived is preferably homologous to or complementary to a sequence which is unique to an HCV genome or the J1 genome. Whether or not a sequence is unique to a genome can be determined by techniques known to those of skill in the art. For example, the sequence can be compared to sequences in databanks, e.g., Genbank, to determine whether it is present in the uninfected host or other organisms. The sequence can also be compared to the known sequences of other viral agents, including those which are known to induce hepatitis, e.g., HAV, HBV, and HDV, and to other members of the Flaviviridae. The correspondence or non-correspondence of the derived sequence to other sequences can also be determined by hybridization under the appropriate stringency conditions. Hybridization techniques for determining the complementarity of nucleic acid sequences are known in the art. See also, for example, Maniatis et al. (1982) MOLECULAR CLONING; A LABORATORY MANUAL (Cold Spring Harbor Press, Cold Spring Harbor, N.Y.). In addition, mismatches of duplex polynucleotides formed by hybridization can be determined by known techniques, including for example, digestion with a nuclease such as S1 that specifically digests single-stranded areas in duplex polynucleotides. Regions from which typical DNA sequences may be derived include, but are not limited to, regions encoding specific epitopes, as well as non-transcribed and/or non-translated regions.

[0030] The J1 polynucleotide is not necessarily physically derived from the nucleotide sequence shown, but may be generated in any manner, including for example, chemical synthesis or DNA replication or reverse transcription or transcription. In addition, combinations of regions corresponding to that of the designated sequence may be modified in ways known in the art to be consistent with an intended use. The polynucleotides may also include one or more labels, which are known to those of skill in the art.

[0031] An amino acid sequence "from" a designated polypeptide or source of polypeptides means that the amino acid sequence is homologous (i.e., identical) to the sequence of the designated polypeptide, or a portion thereof. An amino acid sequence "from" a designated nucleic acid sequence refers to a polypeptide having an amino acid sequence identical to that of a polypeptide encoded in the sequence, or a portion thereof. The J1 amino acid sequences in the polypeptides of the present invention are at least 10 amino acids, more preferably at least about 15 amino acids, and

most preferably at least about 20 amino acids.

[0032] The polypeptides of the present invention are not necessarily translated from a designated nucleic acid sequence; the polypeptides may be generated in any manner, including for example, chemical synthesis, or expression of a recombinant expression system, or isolation from virus. The polypeptides may include one or more analogs of amino acids or unnatural amino acids. Methods of inserting analogs of amino acids into a sequence are known in the art. The polypeptides may also include one or more labels, which are known to those of skill in the art.

[0033] The term "recombinant polynucleotide" as used herein intends a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is linked to a polynucleotide other than that to which it is linked in nature, or (2) does not occur in nature.

[0034] The term "polynucleotide" as used herein refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA, and RNA. It also includes known types of modifications, for example, labels which are known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example proteins (including for e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (e.g., acridine, psoralen, etc.), those containing chelators (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., alpha anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide.

[0035] "Purified polynucleotide" refers to a composition comprising a specified polynucleotide that is substantially free of other components, such composition typically comprising at least about 70% of the specified polynucleotide, more typically at least about 80%, 90% or even 95% to 99% of the specified polynucleotide.

[0036] "Purified polypeptide" refers to a composition comprising a specified polypeptide that is substantially free of other components, such composition typically comprising at least about 70% of the specified polypeptide, more typically at least about 80%, 90% or even 95% to 99% of the specified polypeptide.

[0037] "Recombinant host cells", "host cells", "cells", "cell lines", "cell cultures", and other such terms denote microorganisms or higher eukaryotic cell lines cultured as unicellular entities that can be, or have been, used as recipients for a recombinant vector or other transfer DNA, and include the progeny of the original cell which has been transformed. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation.

[0038] A "replicon" is any genetic element, e.g., a plasmid, a chromosome, a virus, a cosmid, etc. that behaves as an autonomous unit of polynucleotide replication within a cell; i.e., capable of replication under its own control.

[0039] A "cloning vector" is a replicon that can transform a selected host cell and in which another polynucleotide segment is attached, so as to bring about the replication and/or expression of the attached segment. Typically, cloning vectors include plasmids, virus (e.g., bacteriophage vector) and cosmids.

[0040] An "integrating vector" is a vector that does not behave as a replicon in a selected host cell, but has the ability to integrate into a replicon (typically a chromosome) resident in the selected host to stably transform the host.

[0041] An "expression vector" is a construct that can transform a selected host cell and provides for expression of a heterologous coding sequence in the selected host. Expression vectors can be either a cloning vector or an integrating vector.

[0042] A "coding sequence" is a polynucleotide sequence which is transcribed into mRNA and/or translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5'-terminus and a translation stop codon at the 3'-terminus.

A coding sequence can include, but is not limited to mRNA, cDNA, and recombinant polynucleotide sequences.

[0043] "Control sequence" refers to polynucleotide regulatory sequences which are necessary to effect the expression of coding sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism. In prokaryotes, control sequences generally include promoter, ribosomal binding site, and terminators. In eukaryotes generally control sequences include promoters, terminators and, in some instances, enhancers. The term "control sequences" is intended to include, at a minimum, all components the presence of which are necessary for expression, and may also include additional advantageous components.

[0044] "Operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. A control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences.

[0045] An "open reading frame" or ORF is a region of a polynucleotide sequence which encodes a polypeptide; this region may represent a portion of a coding sequence or a total coding sequence.

[0046] "Immunologically cross-reactive" refers to two or more epitopes or polypeptides that are bound by the same

antibody. Cross-reactivity can be determined by any of a number of immunoassay techniques, such as a competition assay.

[0047] As used herein, the term "antibody" refers to a polypeptide or group of polypeptides which comprise at least one epitope. An "antigen binding site" is formed from the folding of the variable domains of an antibody molecule(s) to form three-dimensional binding sites with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows specific binding to form an antibody-antigen complex. An antigen binding site may be formed from a heavy- and/or light-chain domain (VH and VL, respectively), which form hypervariable loops which contribute to antigen binding. The term "antibody" includes, without limitation, chimeric antibodies, altered antibodies, univalent antibodies, Fab proteins, and single-domain antibodies. In many cases, the binding phenomena of antibodies to antigens is equivalent to other ligand/anti-ligand binding.

[0048] As used herein, a "single domain antibody" (dAb) is an antibody which is comprised of an HL domain, which binds specifically with a designated antigen. A dAb does not contain a VL domain, but may contain other antigen binding domains known to exist to antibodies, for example, the kappa and lambda domains. Methods for preparing dAbs are known in the art. See, for example, Ward et al, *Nature* 341: 544 (1989).

[0049] Antibodies may also be comprised of VH and VL domains, as well as other known antigen binding domains. Examples of these types of antibodies and methods for their preparation are known in the art (see, e.g., U.S. Patent No. 4,816,467, which is incorporated herein by reference), and include the following. For example, "vertebrate antibodies" refers to antibodies which are tetramers or aggregates thereof, comprising light and heavy chains which are usually aggregated in a "Y" configuration and which may or may not have covalent linkages between the chains. In vertebrate antibodies, the amino acid sequences of the chains are homologous with those sequences found in antibodies produced in vertebrates, whether in situ or in vitro (for example, in hybridomas). Vertebrate antibodies include, for example, purified polyclonal antibodies and monoclonal antibodies, methods for the preparation of which are described infra.

[0050] "Hybrid antibodies" are antibodies where chains are separately homologous with reference to mammalian antibody chains and represent novel assemblies of them, so that two different antigens are precipitable by the tetramer or aggregate. In hybrid antibodies, one pair of heavy and light chains are homologous to those found in an antibody raised against a first antigen, while a second pair of chains are homologous to those found in an antibody raised against a second antibody. This results in the property of "divalence", i.e., the ability to bind two antigens simultaneously. Such hybrids may also be formed using chimeric chains, as set forth below.

[0051] "Chimeric antibodies" refers to antibodies in which the heavy and/or light chains are fusion proteins. Typically, one portion of the amino acid sequences of the chain is homologous to corresponding sequences in an antibody derived from a particular species or a particular class, while the remaining segment of the chain is homologous to the sequences derived from another species and/or class. Usually, the variable region of both light and heavy chains mimics the variable regions of antibodies derived from one species of vertebrates, while the constant portions are homologous to the sequences in the antibodies derived from another species of vertebrates. However, the definition is not limited to this particular example. Also included is any antibody in which either or both of the heavy or light chains are composed of combinations of sequences mimicking the sequences in antibodies of different sources, whether these sources be from differing classes or different species of origin, and whether or not the fusion point is at the variable/constant boundary. Thus, it is possible to produce antibodies in which neither the constant nor the variable region mimic known antibody sequences. It then becomes possible, for example, to construct antibodies whose variable region has a higher specific affinity for a particular antigen, or whose constant region can elicit enhanced complement fixation, or to make other improvements in properties possessed by a particular constant region.

[0052] Another example is "altered antibodies", which refers to antibodies in which the naturally occurring amino acid sequence in a vertebrate antibody has been varied. Utilizing recombinant DNA techniques, antibodies can be redesigned to obtain desired characteristics. The possible variations are many, and range from the changing of one or more amino acids to the complete redesign of a region, for example, the constant region. Changes in the constant region, in general, to attain desired cellular process characteristics, e.g., changes in complement fixation, interaction with membranes, and other effector functions. Changes in the variable region may be made to alter antigen binding characteristics. The antibody may also be engineered to aid the specific delivery of a molecule or substance to a specific cell or tissue site. The desired alterations may be made by known techniques in molecular biology, e.g., recombinant techniques, site-directed mutagenesis, etc.

[0053] Yet another example are "univalent antibodies", which are aggregates comprised of a heavy-chain/light-chain dimer bound to the Fc (i.e., stem) region of a second heavy chain. This type of antibody escapes antigenic modulation. See, e.g., Glennie et al. *Nature* 295: 712 (1982). Included also within the definition of antibodies are "Fab" fragments of antibodies. The "Fab" region refers to those portions of the heavy and light chains which are roughly equivalent, or analogous, to the sequences which comprise the branch portion of the heavy and light chains, and which have been shown to exhibit immunological binding to a specified antigen, but which lack the effector Fc portion. "Fab" includes aggregates of one heavy and one light chain (commonly known as Fab), as well as tetramers containing the 2H and 2L chains (referred to as F(ab)<sub>2</sub>), which are capable of selectively reacting with a designated antigen or antigen family. Fab anti-

bodies may be divided into subsets analogous to those described above, i.e., "vertebrate Fab", "hybrid Fab", "chimeric Fab", and "altered Fab". Methods of producing Fab fragments of antibodies are known within the art and include, for example, proteolysis, and synthesis by recombinant techniques.

[0054] "Epitope" refers to an antibody binding site usually defined by a polypeptide, but also by non-amino acid hap-  
 5 tens. An epitope could comprise 3 amino acids in a spatial conformation which is unique to the epitope, generally an epitope consists of at least 5 such amino acids, and more usually, consists of at least 8-10 such amino acids.

[0055] "Antigen-antibody complex" refers to the complex formed by an antibody that is specifically bound to an epitope on an antigen.

[0056] "Immunogenic polypeptide" refers to a polypeptide that elicits a cellular and/or humoral immune response in a  
 10 mammal, whether alone or linked to a carrier, in the presence or absence of an adjuvant.

[0057] "Polypeptide" refers to a polymer of amino acids and does not refer to a specific length of the molecule. Thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphoryla-  
 15 tions and the like. Included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

[0058] "Transformation", as used herein, refers to the insertion of an exogenous polynucleotide into a host cell, irre-  
 20 spective of the method used for the insertion, for example, direct uptake, transduction, f-mating or electroporation. The exogenous polynucleotide may be maintained as a non-integrated vector, for example, a plasmid, or alternatively, may be integrated into the host genome.

[0059] A "transformed" host cell refers to both the immediate cell that has undergone transformation and its progeny that maintain the originally exogenous polynucleotide.

[0060] "Treatment" as used herein refers to prophylaxis and/or therapy.

[0061] "Individual", refers to vertebrates, particularly members of the mammalian species, and includes but is not lim-  
 25 ited to domestic animals, sports animals, and primates, including humans.

[0062] "Sense strand" refers to the strand of a double-stranded DNA molecule that is homologous to a mRNA tran-  
 script thereof. The "anti-sense strand" contains a sequence which is complementary to that of the "sense strand".

[0063] "Antibody-containing body component" refers to a component of an individual's body which is a source of the  
 30 antibodies of interest. Antibody-containing body components are known in the art, and include but are not limited to, whole blood and components thereof, plasma, serum, spinal fluid, lymph fluid, the external sections of the respiratory, intestinal, and genitourinary tracts, tears, saliva, milk, white blood cells, and myelomas.

[0064] "Purified HCV" isolate refers to a preparation of HCV particles which has been isolated from the cellular con-  
 35 stituents with which the virus is normally associated, and from other types of viruses which may be present in the infected tissue. The techniques for isolating viruses are known to those of skill in the art, and include, for example, cen-  
 trifugation and affinity chromatography.

[0065] An HCV "particle" is an entire virion, as well as particles which are intermediates in virion formation. HCV par-  
 ticles generally have one or more HCV proteins associated with the HCV nucleic acid.

[0066] "Probe" refers to a polynucleotide which forms a hybrid structure with a sequence in a target polynucleotide,  
 40 due to a complementarity of at least one region in the probe with a region in the target.

[0067] "Biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to,  
 for example, whole blood and components thereof, plasma, serum, spinal fluid, the external sections of the skin, respi-  
 ratory, intestinal, and genitourinary tracts, tears saliva, milk, blood cells, tumors, organs, and also samples of *in vitro* cell  
 culture constituents (including but not limited to conditioned medium resulting from the growth of cells in cell culture  
 medium, putatively virally infected cells, recombinant cells, and cell components).

[0068] The invention pertains to the isolation and characterization of a newly discovered isolate of HCV, J1, its nucle-  
 45 otide sequence, its protein sequence and resulting polynucleotides and polypeptides and antibodies derived therefrom. However, the isolation and characterization of another newly discovered isolate of HCV, J7, is also described herein, although it does not form part of the invention claimed. Isolates J1 and J7 are novel in their nucleotide and amino acid  
 sequences, and are believed to be characteristic of HCV isolates from Japan and other Asian countries.

[0069] The nucleotide sequences derived from HCV J1 are useful as probes to diagnose the presence of virus in sam-  
 50 ples, and to isolate other naturally occurring variants of the virus. These nucleotide sequences also make available polypeptide sequences of HCV antigens encoded within the J1 genome and permits the production of polypeptides  
 which are useful as standards or reagents in diagnostic tests and/or as components or vaccines. Antibodies, both pol-  
 yclonal and monoclonal, directed against HCV epitopes contained within these polypeptide sequences are also useful  
 55 for diagnostic tests, as therapeutic agents, for screening of antiviral agents, and for the isolation of the NANBH virus. In  
 addition, by utilizing probes derived from the sequences disclosed herein it is possible to isolate and sequence other  
 portions of the J1 genome, thus giving rise to additional probes and polypeptides which are useful in the diagnosis  
 and/or treatment, both prophylactic and therapeutic, of NANBH.



[0070] The availability of the HCV/J1 nucleotide sequences enable the construction of polynucleotide probes and polypeptides useful in diagnosing NANBH due to HCV infection and in screening blood donors as well as donated blood and blood products for infection. For example, from the sequences it is possible to synthesize DNA oligomers of about 8-10 nucleotides, or larger, which are useful as hybridization probes to detect the presence of HCV RNA in, for example, sera of subjects suspected of harboring the virus, or for screening donated blood for the presence of the virus. The HCV/J1 sequences also allow the design and production of HCV specific polypeptides which are useful as diagnostic reagents for the presence of antibodies raised during NANBH. Antibodies to purified polypeptides derived from the HCV/J1 sequences may also be used to detect viral antigens in infected individuals and in blood.

[0071] Knowledge of these HCV/J1 sequences also enable the design and production of polypeptides which may be used as vaccines against HCV and also for the production of antibodies, which in turn may be used for protection against the disease, and/or for therapy of HCV infected individuals. Moreover, the disclosed HCV/J1 sequences enable further characterization of the HCV genome. Polynucleotide probes derived from these sequences, as well as from the HCV genome, may be used to screen cDNA libraries for additional viral cDNA sequences, which, in turn, may be used to obtain additional overlapping sequences. See, e.g., EPO Pub. No. 318,216.

[0072] The HCV/J1 polynucleotide sequences, the polypeptides derived therefrom and the antibodies directed against these polypeptides, are useful in the isolation and identification of the BB-NANBV agent(s). For example, antibodies directed against HCV epitopes contained in polypeptides derived from the HCV/J1 sequences may be used in processes based upon affinity chromatography to isolate the virus. Alternatively, the antibodies may be used to identify viral particles isolated by other techniques. The viral antigens and the genomic material within the isolated viral particles may then be further characterized.

[0073] The information obtained from further sequencing of the HCV/J1 genome, as well as from further characterization of the HCV/J1 antigens and characterization of the genomes enable the design and synthesis of additional probes and polypeptides and antibodies which may be used for diagnosis, for prevention, and for therapy of HCV induced NANBH, and for screening for infected blood and blood-related products.

[0074] The availability of HCV/J1 cDNA sequences permits the construction of expression vectors encoding antigenically active regions of the polypeptide encoded in either strand. These antigenically active regions may be derived from coat or envelope antigens or from core antigens, or from antigens which are non-structural including, for example, polynucleotide binding proteins, polynucleotide polymerase(s), and other viral proteins required for the replication and/or assembly of the virus particle. Fragments encoding the desired polypeptides are derived from the cDNA clones using conventional restriction digestion or by synthetic methods, and are ligated into vectors which may, for example, contain portions of fusion sequences such as beta-galactosidase or superoxide dismutase (SOD). Methods and vectors which are useful for the production of polypeptides which contain fusion sequences of SOD are described in EPO Pub. No. 196,056. Vectors encoding fusion polypeptides of SOD and HCV polypeptides are described in EPO Pub. No. 318,216. Any desired portion of the HCV cDNA containing an open reading frame, in either sense strand, can be obtained as a recombinant polypeptide, such as a mature or fusion protein. Alternatively, a polypeptide encoded in the cDNA can be provided by chemical synthesis.

[0075] The DNA encoding the desired polypeptide, whether in fused or mature form, and whether or not containing a signal sequence to permit secretion, may be ligated into expression vectors suitable for any convenient host. Both eukaryotic and prokaryotic host systems are presently used in forming recombinant polypeptides, and a summary of some of the more common control systems and host cell is given below. The polypeptide produced in such host cells is then isolated from lysed cells or from the culture medium and purified to the extent needed for its intended use. Purification may be by techniques known in the art, for example, differential extraction, salt fractionation, chromatography on ion exchange resins, affinity chromatography, centrifugation, and the like. See, for example, Methods in Enzymology for a variety of methods for purifying proteins.

[0076] Such recombinant or synthetic HCV polypeptides can be used as diagnostics, or those which give rise to neutralizing antibodies may be formulated into vaccines. Antibodies raised against these polypeptides can also be used as diagnostics, or for passive immunotherapy. In addition, antibodies to these polypeptides are useful for isolating and identifying HCV particles.

[0077] The HCV antigens may also be isolated from HCV virions. The virions may be grown in HCV infected cells in tissue culture, or in an infected host.

[0078] While the polypeptides of the present invention may comprise a substantially complete viral domain, in many applications all that is required is that the polypeptide comprise an antigenic or immunogenic region of the virus. An antigenic region of a polypeptide is generally relatively small--typically 8 to 10 amino acids or less in length. Fragments of as few as 5 amino acids may characterize an antigenic region. These segments may correspond to regions of HCV/J1 epitopes. Accordingly, using the cDNAs of HCV/J1 as a basis, DNAs encoding short segments of HCV/J1 polypeptides can be expressed recombinantly either as fusion proteins, or as isolated polypeptides. In addition, short amino acid sequences can be conveniently obtained by chemical synthesis.

[0079] In instances wherein the synthesized polypeptide is correctly configured so as to provide the correct epitope,

but is too small to be immunogenic, the polypeptide may be linked to a suitable carrier. A number of techniques for obtaining such linkage are known in the art, including the formation of disulfide linkages using N-succinimidyl-3-(2-pyridyl-thio)propionate (SPDP) and succinimidyl 4-(N-maleimido-methyl)cyclohexane-1-carboxylate (SMCC) obtained from Pierce Company, Rockford, Illinois, (if the peptide lacks a sulfhydryl group, this can be provided by addition of a cysteine residue.) These reagents create a disulfide linkage between themselves and peptide cysteine residues on one protein and an amide linkage through the epsilon-amino on a lysine, or other free amino group in the other. A variety of such disulfide/amide-forming agents are known. See, for example, *Immun. Rev.* (1982) 62:185. Other bifunctional coupling agents form a thioether rather than a disulfide linkage. Many of these thio-ether-forming agents are commercially available and include reactive esters of 6-maleimidocaproic acid, 2-bromoacetic acid, 2-iodoacetic acid, 4-(N-maleimido-methyl)cyclohexane-1-carboxylic acid, and the like. The carboxyl groups can be activated by combining them with succinimide or 1-hydroxyl-2-nitro-4-sulfonic acid, sodium salt. Additional methods of coupling antigens employ the rotavirus/"binding peptide" system described in EPO Pub. No. 259,149, the disclosure of which is incorporated herein by reference. The foregoing list is not meant to be exhaustive, and modifications of the named compounds can clearly be used.

[0080] Any carrier may be used which does not itself induce the production of antibodies harmful to the host. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins; polysaccharides, such as latex functionalized sepharose, agarose, cellulose, cellulose beads and the like; polymeric amino acids, such as polyglutamic acid, polylysine, and the like; amino acid copolymers; and inactive virus particles. Especially useful protein substrates are serum albumins, keyhole limpet hemocyanin, immunoglobulin molecules, thyroglobulin, ovalbumin, tetanus toxoid, and other proteins well known to those skilled in the art.

[0081] In addition to full-length viral proteins, polypeptides comprising truncated HCV amino acid sequences encoding at least one viral epitope are useful immunological reagents. For example, polypeptides comprising such truncated sequences can be used as reagents in an immunoassay. These polypeptides also are candidate subunit antigens in compositions for antiserum production or vaccines. While these truncated sequences can be produced by various known treatments of native viral protein, it is generally preferred to make synthetic or recombinant polypeptides comprising an HCV sequence. Polypeptides comprising these truncated HCV sequences can be made up entirely of HCV sequences (one or more epitopes, either contiguous or noncontiguous), or HCV sequences and heterologous sequences in a fusion protein. Useful heterologous sequences include sequences that provide for secretion from a recombinant host, enhance the immunological reactivity of the HCV epitope(s), or facilitate the coupling of the polypeptide to an immunoassay support or a vaccine carrier. See, e.g., EPO Pub. No. 116,201; U.S. Pat. No. 4,722,840; EPO Pub. No. 259,149; U.S. Pat. No. 4,629,783, the disclosures of which are incorporated herein by reference.

[0082] The size of polypeptides comprising the truncated HCV sequences can vary widely, the minimum size being a sequence of sufficient size to provide an HCV epitope, while the maximum size is not critical. In some applications, the maximum size usually is not substantially greater than that required to provide the desired HCV epitopes and function(s) of the heterologous sequence, if any. Typically, the truncated HCV amino acid sequence will range from about 5 to about 100 amino acids in length. More typically, however, the HCV sequence will be a maximum of about 50 amino acids in length, preferably a maximum of about 30 amino acids. It is usually desirable to select HCV sequences of at least about 10, 12 or 15 amino acids, up to a maximum of about 20 or 25 amino acids.

[0083] Truncated HCV amino acid sequences comprising epitopes can be identified in a number of ways. For example, the entire viral protein sequence can be screened by preparing a series of short peptides that together span the entire protein sequence. By starting with, for example, 100-mer polypeptides, it would be routine to test each polypeptide for the presence of epitope(s) showing a desired reactivity, and then testing progressively smaller and overlapping fragments from an identified 100-mer to map the epitope of interest. Screening such peptides in an immunoassay is within the skill of the art. It is also known to carry out a computer analysis of a protein sequence to identify potential epitopes, and then prepare oligopeptides comprising the identified regions for screening. It is appreciated by those of skill in the art that such computer analysis of antigenicity does not always identify an epitope that actually exists, and can also incorrectly identify a region of the protein as containing an epitope.

[0084] The observed relationship of the putative polyproteins of HCV and the Flaviviruses allows a prediction of the putative domains of the HCV "non-structural" (NS) proteins. The locations of the individual NS proteins in the putative Flavivirus precursor polypeptide are fairly well-known. Moreover, these also coincide with observed gross fluctuations in the hydrophobicity profile of the polypeptide. It is established that NS5 of Flaviviruses encodes the virion polymerase, and that NS1 corresponds with a complement fixation antigen which has been shown to be an effective vaccine in animals. Recently, it has been shown that a flaviviral protease function resides in NS3. Due to the observed similarities between HCV and the Flaviviruses, deductions concerning the approximate locations of the corresponding protein domains and functions in the HCV polypeptide are possible. Figure 11 is a schematic of putative domains of the HCV polypeptide. The expression of polypeptides containing these domains in a variety of recombinant host cells, including, for example, bacteria, yeast, insect, and vertebrate cells, should give rise to important immunological reagents which can be used for diagnosis, detection, and vaccines.

[0085] Although the non-structural protein region of the putative polyproteins of the HCV isolate described herein and of Flaviviruses appears to be generally similar, there is less similarity between the putative structural regions which are towards the N-terminus. In this region, there is a greater divergence in sequence, and in addition, the hydrophobic profile of the two regions show less similarity. This "divergence" begins in the N-terminal region of the putative NS1 domain in HCV, and extends to the presumed N-terminus. Nevertheless, it is still possible to predict the approximate locations of the putative nucleocapsid (N-terminal basic domain) and E (generally hydrophobic) domains within the HCV polyprotein. From these predictions it may be possible to identify approximate regions of the HCV polyprotein that could correspond with useful immunological reagents. For example, the E and NS1 proteins of Flaviviruses are known to have efficacy as protective vaccines. These regions, as well as some which are shown to be antigenic in the HCV1, for example those within putative NS3, C, and NS5, etc., should also provide diagnostic reagents.

[0086] The immunogenicity of the HCV sequences may also be enhanced by preparing the sequences fused to or assembled with particle-forming proteins such as, for example, hepatitis B surface antigen or rotavirus VP6 antigen. Constructs wherein the HCV epitope is linked directly to the particle-forming protein coding sequences produce hybrids which are immunogenic with respect to the HCV epitope. In addition, all of the vectors prepared include epitopes specific to HBV, having various degrees of immunogenicity, such as, for example, the pre-S peptide. Thus, particles constructed from particle forming protein which include HCV sequences are immunogenic with respect to HCV and particle-form protein. See, e.g., U.S. Pat. No. 4,722,840; EPO Pub No. 175,261; EPO Pub. No. 259,149; Michelle et al. (1984) Int. Symposium on Viral Hepatitis.

[0087] Vaccines may be prepared from one or more immunogenic polypeptides derived from HCV/J1. The observed homology between HCV and Flaviviruses provides information concerning the polypeptides which are likely to be most effective as vaccines, as well as the regions of the genome in which they are encoded. The general structure of the Flavivirus genome is discussed in Rice et al. (1986) in THE VIRUSES: THE TOGAVIRIDAE AND FLAVIVIRIDAE (Series eds. Fraenkel-Conrat and Wagner, Vol eds. Schlesinger and Schlesinger, Plenum Press). The flavivirus genomic RNA is believed to be the only virus-specific mRNA species, and it is translated into the three viral structural proteins, i.e., C, M, and E, as well as two large nonstructural proteins, NS4 and NS5, and a complex set of smaller nonstructural proteins. It is known that major neutralizing epitopes for Flaviviruses reside in the E (envelope) protein. Roehrig (1986) in THE VIRUSES: THE TOGAVIRIDAE AND FLAVIVIRIDAE (Series eds. Fraenkel-Conrat and Wagner, Vol eds. Schlesinger and Schlesinger, Plenum Press). The corresponding HCV E gene and polypeptide encoding region may be predicted, based upon the homology to Flaviviruses. Thus, vaccines may be comprised of recombinant polypeptides containing epitopes of HCV E. These polypeptides may be expressed in bacteria, yeast, or mammalian cells, or alternatively may be isolated from viral preparations. It is also anticipated that the other structural proteins may also contain epitopes which give rise to protective anti-HCV antibodies. Thus, polypeptides containing the epitopes of E, C, and M may also be used, whether singly or in combination, in HCV vaccines.

[0088] In addition to the above, it has been shown that immunization with NS1 (nonstructural protein 1), results in protection against yellow fever. Schlesinger et al (1986) J. Virol. 60:1153. This is true even though the immunization does not give rise to neutralizing antibodies. Thus, particularly since this protein appears to be highly conserved among Flaviviruses, it is likely that HCV NS1 will also be protective against HCV infection. Moreover, it also shows that nonstructural proteins may provide protection against viral pathogenicity, even if they do not cause the production of neutralizing antibodies.

[0089] In view of the above, multivalent vaccines against HCV may be comprised of one or more epitopes from one or more structural proteins, and/or one or more epitopes from one or more nonstructural proteins. These vaccines may be comprised of, for example, recombinant HCV polypeptides and/or polypeptides isolated from the virions. In particular, vaccines are contemplated comprising one or more of the following HCV proteins, or subunit antigens derived therefrom: E, NS1, C, NS2, NS3, NS4 and NS5. Particularly preferred are vaccines comprising E and/or NS1, or subunits thereof. In addition, it may be possible to use inactivated HCV in vaccines; inactivation may be by the preparation of viral lysates, or by other means known in the art to cause inactivation of Flaviviruses, for example, treatment with organic solvents or detergents, or treatment with formalin. Moreover, vaccines may also be prepared from attenuated HCV strains or from hybrid viruses such as vaccinia vectors known in the art [Brown et al. Nature 319: 549-550 (1986)].

[0090] The preparation of vaccines which contain immunogenic polypeptide(s) as active ingredients is known to one skilled in the art. Typically, such vaccines are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. The preparation may also be emulsified, or the protein encapsulated in liposomes. The active immunogenic ingredients are often mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the vaccine may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the vaccine. Examples of adjuvants which may be effective include but are not limited to: aluminum hydroxide, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-

alanine-2-(1'-2'-dipalmitoyl-sn-glycerol-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween 80 emulsion. The effectiveness of an adjuvant may be determined by measuring the amount of antibodies directed against an immunogenic polypeptide containing an HCV antigenic sequence resulting from administration of this polypeptide in vaccines which are also comprised of the various adjuvants.

[0091] The vaccines are conventionally administered parenterally, by injection, usually, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration include suppositories and, in some cases, oral formulations. For suppositories, traditional binders and carriers may include, for example, polyalkylene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1%-2%. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain 10%-95% of active ingredient, preferably 25%-70%.

[0092] The proteins may be formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts include the acid addition salts (formed with free amino groups of the peptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids such as acetic, oxalic, tartaric, maleic, and the like. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

[0093] The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be prophylactically and/or therapeutically effective. The quantity to be administered, which is generally in the range of 5 micrograms to 250 micrograms of antigen per dose, depends on the subject to be treated, capacity of the subject's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered may depend on the judgment of the practitioner and may be peculiar to each subject.

[0094] The vaccine may be given in a single dose schedule, or preferably in a multiple dose schedule. A multiple dose schedule is one in which a primary course of vaccination may be with 1-10 separate doses, followed by other doses given at subsequent time intervals required to maintain and/or reinforce the immune response, for example, at 1-4 months for a second dose, and if needed, a subsequent dose(s) after several months. The dosage regimen will also, at least in part, be determined by the need of the individual and be dependent upon the judgment of the practitioner.

[0095] In addition, the vaccine containing the immunogenic HCV antigen(s) may be administered in conjunction with other immunoregulatory agents, for example, immune globulins.

[0096] The immunogenic polypeptides prepared as described above are used to produce antibodies, both polyclonal and monoclonal. If polyclonal antibodies are desired, a selected mammal (e.g., mouse, rabbit, goat, horse, etc.) is immunized with an immunogenic polypeptide bearing an HCV epitope(s). Serum from the immunized animal is collected and treated according to known procedures. If serum containing polyclonal antibodies to an HCV epitope contains antibodies to other antigens, the polyclonal antibodies can be purified by immunoaffinity chromatography. Techniques for producing and processing polyclonal antisera are known in the art, see for example, Mayer and Walker, eds. (1987) IMMUNOCHEMICAL METHODS IN CELL AND MOLECULAR BIOLOGY (Academic Press, London).

[0097] Monoclonal antibodies directed against HCV epitopes can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by hybridomas is well known. Immortal antibody-producing cell lines can be created by cell fusion, and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. See, e.g., M. Schreier et al. (1980) HYBRIDOMA TECHNIQUES; Hammerling et al. (1981), MONOCLONAL ANTIBODIES AND T-CELL HYBRIDOMAS; Kennett et al. (1980) MONOCLONAL ANTIBODIES; see also, U.S. Patent Nos. 4,341,761; 4,399,121; 4,427,783; 4,444,887; 4,466,917; 4,472,500; 4,491,632; and 4,493,890. Panels of monoclonal antibodies produced against HCV epitopes can be screened for various properties; i.e., for isotype, epitope affinity, etc.

[0098] Antibodies, both monoclonal and polyclonal, which are directed against HCV epitopes are particularly useful in diagnosis, and those which are neutralizing are useful in passive immunotherapy. Monoclonal antibodies, in particular, may be used to raise anti-idiotype antibodies.

[0099] Anti-idiotype antibodies are immunoglobulins which carry an "internal image" of the antigen of the infectious agent against which protection is desired. Techniques for raising anti-idiotype antibodies are known in the art. See, e.g., Grzych (1985), Nature 316:74; MacNamara et al. (1984), Science 226:1325; Uytdehaag et al (1985), J. Immunol. 134:1225. These anti-idiotype antibodies may also be useful for treatment and/or diagnosis of NANBH, as well as for an elucidation of the immunogenic regions of HCV antigens.

[0100] Using the HCV/J1 polynucleotide sequences as a basis, oligomers of approximately 8 nucleotides or more can be prepared, either by excision or synthetically, which hybridize with the HCV genome and are useful in identification of the viral agent(s), further characterization of the viral genome(s), as well as in detection of the virus(es) in diseased indi-

viduals. The probes for HCV polynucleotides (natural or derived) are a length which allows the detection of unique viral sequences by hybridization. While 6-8 nucleotides may be a workable length, sequences of about 10-12 nucleotides are preferred, and about 20 nucleotides appears optimal. These probes can be prepared using routine methods, including automated oligonucleotide synthetic methods. Among useful probes, for example, are the clones disclosed herein, as well as the various oligomers useful in probing cDNA libraries, set forth below. A complement to any unique portion of the HCV genome will be satisfactory. For use as probes, complete complementarity is desirable, though it may be unnecessary as the length of the fragment is increased.

[0101] For use of such probes as diagnostics, the biological sample to be analyzed, such as blood or serum, may be treated, if desired, to extract the nucleic acids contained therein. The resulting nucleic acid from the sample may be subjected to gel electrophoresis or other size separation techniques; alternatively, the nucleic acid sample may be dot blotted without size separation. The probes are then labeled. Suitable labels, and methods for labeling probes are known in the art, and include, for example, radioactive labels incorporated by nick translation or kinasing, biotin, fluorescent probes, and chemiluminescent probes. The nucleic acids extracted from the sample are then treated with the labeled probe under hybridization conditions of suitable stringencies. Usually high stringency conditions are desirable in order to prevent false positives. The stringency of hybridization is determined by a number of factors during hybridization and during the washing procedure, including temperature, ionic strength, length of time, and concentration of formamide. These factors are outlined in, for example, Maniatis, T. (1982) MOLECULAR CLONING; A LABORATORY MANUAL (Cold Spring Harbor Press, Cold Spring Harbor, N.Y.).

[0102] Generally, it is expected that the HCV genome sequences will be present in serum of infected individuals at relatively low levels, i.e., at approximately  $10^2$ - $10^3$  chimp infectious doses (CID) per ml. This level may require that amplification techniques be used in hybridization assays. Such techniques are known in the art. For example, the Enzo Biochemical Corporation "Bio-Bridge" system uses terminal deoxynucleotide transferase to add unmodified 3'-poly-dT-tails to a DNA probe. The poly dT-tailed probe is hybridized to the target nucleotide sequence, and then to a biotin-modified poly-A. PCT App. No. 84/03520 and EPO Pub. No. 124,221 describe a DNA hybridization assay in which: (1) analyte is annealed to a single-stranded DNA probe that is complementary to an enzyme-labeled oligonucleotide; and (2) the resulting tailed duplex is hybridized to an enzyme-labeled oligonucleotide. EPO Pub. No. 204,510 describes a DNA hybridization assay in which analyte DNA is contacted with a probe that has a tail, such as a poly-dT tail, an amplifier strand that has a sequence that hybridizes to the tail of the probe, such as a poly-A sequence, and which is capable of binding a plurality of labeled strands.

[0103] A particularly desirable technique may first involve amplification of the target HCV sequences in sera approximately 10,000-fold, i.e., to approximately  $10^6$  sequences/ml. This may be accomplished, for example, by the polymerase chain reactions (PCR) technique described which is by Saiki et al. (1986) Nature 324:163, Mullis, U.S. Patent No. 4,683,195, and Mullis et al. U.S. Patent No. 4,683,202. The amplified sequence(s) may then be detected using a hybridization assay which is described in co-pending European Publication No. 317-077 and Japanese application No. 63-260347, which are assigned to the herein assignee, and are hereby incorporated herein by reference. These hybridization assays, which should detect sequences at the level of  $10^6$ /ml, utilize nucleic acid multimers which bind to single-stranded analyte nucleic acid, and which also bind to a multiplicity of single-stranded labeled oligonucleotides. A suitable solution phase sandwich assay which may be used with labeled polynucleotide probes, and the methods for the preparation of probes is described in EPO Pub. No. 225,807 which is hereby incorporated herein by reference.

[0104] The probes can be packaged into diagnostic kits. Diagnostic kits include the probe DNA, which may be labeled; alternatively, the probe DNA may be unlabeled and the ingredients for labeling may be included in the kit in separate containers. The kit may also contain other suitably packaged reagents and materials needed for the particular hybridization protocol, for example, standards, wash buffers, as well as instructions for conducting the test.

[0105] Both the HCV/J1 polypeptides which react immunologically with serum containing HCV antibodies and the antibodies raised against the HCV specific epitopes in these polypeptides are useful in immunoassays to detect presence of HCV antibodies, or the presence of the virus and/or viral antigens, in biological samples. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. An immunoassay for anti-HCV antibody may utilize one viral epitope or several viral epitopes. When multiple epitopes are used, the epitopes may be derived from the same or different viral polypeptides, and may be in separate recombinant or natural polypeptides, or together in the same recombinant polypeptides.

[0106] An immunoassay for viral antigen may use, for example, a monoclonal antibody directed towards a viral epitope, a combination of monoclonal antibodies directed towards epitopes of one viral polypeptide, monoclonal antibodies directed towards epitopes of different viral polypeptides, polyclonal antibodies directed towards the same viral antigen, polyclonal antibodies directed towards different viral antigens or a combination of monoclonal and polyclonal antibodies.

[0107] Immunoassay protocols may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide. The labels may be, for example, fluorescent, chemiluminescent, radioactive,

or dye molecules. Assays which amplify the signals from the probe are also known. Examples of which are assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

[0108] Typically, an immunoassay for anti-HCV antibody will involve selecting and preparing the test sample, such as a biological sample, and then incubating it with an antigenic (i.e., epitope-containing) HCV polypeptide under conditions that allow antigen-antibody complexes to form. Such conditions are well known in the art. In a heterogeneous format, the polypeptide is bound to a solid support to facilitate separation of the sample from the polypeptide after incubation. Examples of solid supports that can be used are nitrocellulose, in membrane or microtiter well form, polyvinylchloride, in sheets or microtiter wells, polystyrene latex, in beads or microtiter plates, polyvinylidene fluoride, known as Immobilon™, diazotized paper, nylon membranes, activated beads, and Protein A beads. Most preferably, the Dynatech, Immulon™ 1 microtiter plate or the 0.25-inch polystyrene beads, which Spec finished by Precision Plastic Ball, are used in the heterogeneous format. The solid support is typically washed after separating it from the test sample. In a homogeneous format, the test sample is incubated with antigen in solution, under conditions that will precipitate any antigen-antibody complexes that are formed, as is known in the art. The precipitated complexes are then separated from the test sample, for example, by centrifugation. The complexes formed comprising anti-HCV antibody are then detected by any of a number of techniques. Depending on the format, the complexes can be detected with labeled anti-xenogeneic Ig or, if a competitive format is used, by measuring the amount of bound, labeled competing antibody.

[0109] In immunoassays where HCV polypeptides are the analyte, the test sample, typically a biological sample, is incubated with anti-HCV antibodies again under conditions that allow the formation of antigen-antibody complexes. Various formats can be employed, such as a "sandwich" assay where antibody bound to a solid support is incubated with the test sample; washed; incubated with a second, labeled antibody to the analyte; and the support is washed again. Analyte is detected by determining if the second antibody is bound to the support. In a competitive format, which can be either heterogeneous or homogeneous, a test sample is usually incubated with an antibody and a labeled, competing antigen either sequentially or simultaneously. These and other formats are well known in the art.

[0110] The Flavivirus model for HCV allows predictions regarding the likely location of diagnostic epitopes for the virion structural proteins. The C, pre-M, M, and E domains are all likely to contain epitopes of significant potential for detecting viral antigens, and particularly for diagnosis. Similarly, domains of the nonstructural proteins are expected to contain important diagnostic epitopes (e.g., NS5 encoding a putative polymerase; and NS1 encoding a putative complement-binding antigen). Recombinant polypeptides, or viral polypeptides, which include epitopes from these specific domains may be useful for the detection of viral antibodies in infections blood donors and infected patients. In addition, antibodies directed against the E and/or M proteins can be used in immunoassays for the detection of viral antigens in patients with HCV caused NANBH, and in infectious blood donors. Moreover, these antibodies may be extremely useful in detecting acute-phase donors and patients.

[0111] Antigenic regions of the putative polyprotein can be mapped and identified by screening the antigenicity of bacterial expression products of HCV cDNAs which encode portions of the polyprotein. Other antigenic regions of HCV may be detected by expressing the portions of the HCV cDNAs in other expression systems, including yeast systems and cellular systems derived from insects and vertebrates. In addition, studies giving rise to an antigenicity index and hydrophobicity/hydrophilicity profile give rise to information concerning the probability of a region's antigenicity. Efficient detection systems may include the use of panels of epitopes. The epitopes in the panel may be constructed into one or multiple polypeptides.

[0112] Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the polypeptides of the invention containing HCV epitopes or antibodies directed against HCV epitopes in suitable containers, along with the remaining reagents and materials required for the conduct of the assay (e.g., wash buffers, detection means like labeled anti-human Ig, labeled anti-HCV, or labeled HCV antigen), as well as a suitable set of assay instructions.

[0113] The HCV/J1 nucleotide sequence information described herein may be used to gain further information on the sequence of the HCV genomes, and for identification and isolation of additional HCV isolates related to J1. This information, in turn, can lead to additional polynucleotide probes, polypeptides derived from the HCV genome, and antibodies directed against HCV epitopes which would be useful for the diagnosis and/or treatment of HCV caused NANBH.

[0114] The HCV/J1 nucleotide sequence information herein is useful for the design of probes for the isolation of additional sequences which are derived from as yet undefined regions of the HCV genomes from which the J1 sequences are derived. For example, labeled probes containing a sequence of approximately 8 or more nucleotides; and preferably 20 or more nucleotides, which are derived from regions close to the 5'-termini or 3'-termini of the family of HCV cDNA sequences disclosed in the examples may be used to isolate overlapping cDNA sequences from HCV cDNA libraries. These sequences which overlap the cDNAs in the above-mentioned clones, but which also contain sequences derived from regions of the genome from which the cDNA in the above mentioned clones are not derived, may then be used to synthesize probes for identification of other overlapping fragments which do not necessarily overlap the cDNAs described below. Methods for constructing cDNA libraries are known in the art. See, e.g. EPO Pub. No. 318,216. It is particularly preferred to prepare libraries from the serum of Japanese and other Asian patients diagnosed as having



NANBH demonstrating antibody to HCV1 antigens; these are believed to be the most likely candidates for carriers of HCV/J1 or related isolates.

[0115] HCV particles may be isolated from the sera from individuals with NANBH or from cell cultures by any of the methods known in the art, including for example, techniques based on size discrimination such as sedimentation or exclusion methods, or techniques based on density such as ultracentrifugation in density gradients, or precipitation with agents such as polyethylene glycol, or chromatography on a variety of materials such as anionic or cationic exchange materials, and materials which bind due to hydrophobicity.

[0116] A preferred method of isolating HCV particles or antigen is by immunoaffinity columns. Techniques for immunoaffinity chromatography are known in the art, including techniques for affixing antibodies to solid supports so that they retain their immunoselective activity. The techniques may be those in which the antibodies are adsorbed to the support (see, for example, Kurstak in ENZYME IMMUNODIAGNOSIS, page 31-37), as well as those in which the antibodies are covalently linked to the support. Generally, the techniques are similar to those used for covalent linking of antigens to a solid support, described above. However, spacer groups may be included in the bifunctional coupling agents so that the antigen binding site of the antibody remains accessible. The antibodies may be monoclonal, or polyclonal, and it may be desirable to purify the antibodies before their use in the immunoassay.

[0117] The general techniques used in extracting the genome from a virus, preparing and probing a cDNA library, sequencing clones, constructing expression vectors, transforming cells, performing immunological assays such as radioimmunoassays and ELISA assays, for growing cells in culture, and the like are known in the art and laboratory manuals are available describing these techniques. However, as a general guide, the following sets forth some sources currently available for such procedures, and for materials useful in carrying them out.

[0118] Both prokaryotic and eukaryotic host cells may be used for expression of desired coding sequences when appropriate control sequences which are compatible with the designated host are used. Among prokaryotic hosts, *E. coli* is most frequently used. Expression control sequences for prokaryotes include promoters, optionally containing operator portions, and ribosome binding sites. Transfer vectors compatible with prokaryotic hosts are commonly derived from, for example, pBR322, a plasmid containing operons conferring ampicillin and tetracycline resistance, and the various pUC vectors, which also contain sequences conferring antibiotic resistance markers. These markers may be used to obtain successful transformants by selection. Commonly used prokaryotic control sequences include the Beta-lactamase (penicillinase) and lactose promoter systems (Chang et al. (1977), Nature 198:1056, the tryptophan (trp) promoter system (Goeddel et al. (1980) Nucleic Acid Res. 8:4057), and the lambda-derived P<sub>L</sub> promoter and N gene ribosome binding site (Shimatake et al. (1981) Nature 292:128) and the hybrid lac promoter (De Boer et al. (1983) Proc. Natl. Acad. Sci. USA 292:128) derived from sequences of the trp and lac UV5 promoters. The foregoing systems are particularly compatible with *E. coli*; if desired, other prokaryotic hosts such as strains of *Bacillus* or *Pseudomonas* may be used, with corresponding control sequences.

[0119] Eukaryotic hosts include yeast and mammalian cells in culture systems. *Saccharomyces cerevisiae*, *Saccharomyces carlsbergensis*, *Klebsiella lactis* and *Pichia pastoris* are the most commonly used yeast hosts, and are convenient fungal hosts. Yeast compatible vectors carry markers which permit selection of successful transformants by conferring prototrophy to auxotrophic mutants or resistance to heavy metals on wild-type strains. Yeast compatible vectors may employ the 2 micron origin of replication (Broach et al. (1983) Math. Enz. 101:307), the combination of CEN3 and ARS1 or other means for assuring replication, such as sequences which will result in incorporation of an appropriate fragment into the host cell genome. Control sequences for yeast vectors are known in the art and include promoters for the synthesis of glycolytic enzymes (Hess et al. (1968) J. Adv. Enzyme Eng. 7:149; Holland et al. (1978), J. Biol. Chem. 255:1385), including the promoter for 3-phosphoglycerate kinase (Hitzeman (1980), J. Biol. Chem. 255:2073). Terminators may also be included, such as those derived from the enolase gene (Holland (1981), J. Biol. Chem. 256:1385). Particularly useful control systems are those which comprise the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter or alcohol dehydrogenase (ADH) regulatable promoter, terminators also derived from GAPDH, and if secretion is desired, leader sequence from yeast alpha factor. In addition, the transcriptional regulatory region and the transcriptional initiation region which are operably linked may be such that they are not naturally associated in the wild-type organism. These systems are described in detail in EPO Pub. No. 120,551; EPO Pub. No. 116,201; and EPO Pub. No. 164,556 all of which are incorporated herein by reference.

[0120] Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including HeLa cells, Chinese hamster ovary (CHO) cells, baby hamster kidney (BHK) cells, and a number of other cell lines. Suitable promoters for mammalian cells are also known in the art and include viral promoters such as that from Simian Virus 40 (SV40) (Fiers (1978), Nature 273:113), Rous sarcoma virus (RSV), adenovirus (ADV), and bovine papilloma virus (BPV). Mammalian cells may also require terminator sequences and poly A addition sequences; enhancer sequences which increase expression may also be included, and sequences which cause amplification of the gene may also be desirable. These sequences are known in the art. Vectors suitable for replication in mammalian cells may include viral replicons, or sequences which insure integration of the appropriate sequences encoding NANBV epitopes into the host genome.

[0121] The vaccinia virus system can also be used to express foreign DNA in mammalian cells. To express heterologous genes, the foreign DNA is usually inserted into the thymidine kinase gene of the vaccinia virus and then infected cells can be selected. This procedure is known in the art and further information can be found in these references [Mackett et al. J. Virol. 49: 857-864 (1984) and Chapter 7 in DNA Cloning, Vol. 2, IRL Press].

5 [0122] In addition, viral antigens can be expressed in insect cells by the Baculovirus system. A general guide to baculovirus expression by Summer and Smith is A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures (Texas Agricultural Experiment Station Bulletin No. 1555). To incorporate the heterologous gene into the Baculovirus genome the gene is first cloned into a transfer vector containing some Baculovirus sequences. This transfer vector, when it is cotransfected with wild-type virus into insect cells, will recombine with the wild-type virus. Usually,  
10 the transfer vector will be engineered so that the heterologous gene will disrupt the wild-type Baculovirus polyhedron gene. This disruption enables easy selection of the recombinant virus since the cells infected with the recombinant virus will appear phenotypically different from the cells infected with the wild-type virus. The purified recombinant virus can be used to infect cells to express the heterologous gene. The foreign protein can be secreted into the medium if a signal peptide is linked in frame to the heterologous gene; otherwise, the protein will be bound in the cell lysates. For further  
15 information, see Smith et al Mol. & Cell. Biol. 3:2156-2165 (1983) or Luckow and Summers in Virology 17: 31-39 (1989).

[0123] Transformation may be by any known method for introducing polynucleotides into a host cell, including, for example packaging the polynucleotide in a virus and transducing a host cell with the virus, and by direct uptake of the polynucleotide. The transformation procedure used depends upon the host to be transformed. Bacterial transformation by direct uptake generally employs treatment with calcium or rubidium chloride (Cohen (1972), Proc. Natl. Acad. Sci.  
20 USA 69:2110; Maniatis et al. (1982), MOLECULAR CLONING; A LABORATORY MANUAL (Cold Spring Harbor Press, Cold Spring Harbor, N.Y.). Yeast transformation by direct uptake may be carried out using the method of Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75: 1929. Mammalian transformations by direct uptake may be conducted using the calcium phosphate precipitation method of Graham and Van der Eb (1978), Virology 52:546 or the various known modifications thereof.

25 [0124] Vector construction employs techniques which are known in the art. Site-specific DNA cleavage is performed by treating with suitable restriction enzymes under conditions which generally are specified by the manufacturer of these commercially available enzymes. The cleaved fragments may be separated using polyacrylamide or agarose gel electrophoresis techniques, according to the general procedures found in Methods in Enzymology (1980) 65:499-560. Sticky ended cleavage fragments may be blunt ended using *E. coli* DNA polymerase I (Klenow) in the presence of the  
30 appropriate deoxynucleotide triphosphates (dNTPs) present in the mixture. Treatment with S1 nuclease may also be used, resulting in the hydrolysis of any single stranded DNA portions.

[0125] Ligations are carried out using standard buffer and temperature conditions using T4 DNA ligase and ATP; sticky end ligations require less ATP and less ligase than blunt end ligations. When vector fragments are used as part of a ligation mixture, the vector fragment is often treated with bacterial alkaline phosphatase (BAP) or calf intestinal  
35 alkaline phosphatase to remove the 5'-phosphate and thus prevent religation of the vector; alternatively, restriction enzyme digestion of unwanted fragments can be used to prevent ligation. Ligation mixtures are transformed into suitable cloning hosts, such as *E. coli*, and successful transformants selected by, for example, antibiotic resistance, and screened for the correct construction.

[0126] Synthetic oligonucleotides may be prepared using an automated oligonucleotide synthesizer as described by Warner (1984), DNA 3:401. If desired, the synthetic strands may be labeled with <sup>32</sup>P by treatment with polynucleotide  
40 kinase in the presence of <sup>32</sup>P-ATP, using standard conditions for the reaction. DNA sequences, including those isolated from cDNA libraries, may be modified by known techniques, including, for example site directed mutagenesis, as described by Zoller (1982), Nucleic Acids Res. 10:6487.

[0127] DNA libraries may be probed using the procedure of Grunstein and Hogness (1975), Proc. Natl. Acad. Sci.  
45 USA 73:3961. Briefly, in this procedure, the DNA to be probed is immobilized on nitrocellulose filters, denatured, and prehybridized with a buffer. The percentage of formamide in the buffer, as well as the time and temperature conditions of the prehybridization and subsequent hybridization steps depends on the stringency required. Oligomeric probes which require lower stringency conditions are generally used with low percentages of formamide, lower temperatures,  
50 and longer hybridization times. Probes containing more than 30 or 40 nucleotides such as those derived from cDNA or genomic sequences generally employ higher temperatures, e.g., about 40-42°C, and a high percentage, e.g., 50%, formamide. Following prehybridization, 5'-<sup>32</sup>P-labeled oligonucleotide probe is added to the buffer, and the filters are incubated in this mixture under hybridization conditions. After washing, the treated filters are subjected to autoradiography to show the location of the hybridized probe; DNA in corresponding locations on the original agar plates is used as the source of the desired DNA.

55 [0128] An enzyme-linked immunosorbent assay (ELISA) can be used to measure either antigen or antibody concentrations. This method depends upon conjugation of an enzyme to either an antigen or an antibody, and uses the bound enzyme activity as a quantitative label. To measure antibody, the known antigen is fixed to a solid phase (e.g., a microplate or plastic cup), incubated with test serum dilutions, washed, incubated with anti-immunoglobulin labeled with an



# EP 0 419 182 B1

enzyme, and washed again. Enzymes suitable for labeling are known in the art, and include, for example, horseradish peroxidase. Enzyme activity bound to the solid phase is measured by adding the specific substrate, and determining product formation or substrate utilization colorimetrically. The enzyme activity bound is a direct function of the amount of antibody bound.

- 5 [0129] To measure antigen, a known specific antibody is fixed to the solid phase, the test material containing antigen is added, after an incubation the solid phase is washed, and a second enzyme-labeled antibody is added. After washing, substrate is added, and enzyme activity is estimated colorimetrically, and related to antigen concentration.

## Examples

10

1

[0130] This example describes the cloning of the HCV/J1 and HCV/J7 nucleotide sequences.

- 15 [0131] Both blood samples which were used as a source of HCV virions were found to be positive in an anti-HCV antibody assay. The HCV isolates from these samples were named HCV/J1 and HCV/J7. The infectivity of the blood sample containing the J1 isolate was confirmed by a prospective study of blood transfusion recipients. Dr. Tohru Katayama from the Department of Surgery at the National Tokyo Chest Hospital collected blood from patients who have contracted post-transfusion non-A, non-B hepatitis. He also collected blood samples from the respective blood donors of these patients. Next, these samples were assayed for antibodies to the C100-3 HCV1 antigen (EPO Pub. No. 20 318,216), and blood from one of the donors was found to be positive.

- [0132] Isolation of the RNA from the blood samples began by pelleting virions in the blood sample by ultracentrifugation [Bradley, D.W., McCaustland, K.A., Cook E.H., Schable, C.A., Ebert, J.W. and Maynard, J.E. (1985) Gastroenterology 88, 773-779]. RNA was then extracted from the pellet by the guanidinium/cesium chloride method [Maniatis T., Fritsch, E.F., and Sambrook J. (1982) "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory, Cold 25 Spring Harbor] and further purified by phenol/chloroform extraction in the presence of urea, [Berk, A.J. Lee, F., Harrison, T., Williams, J. and Sharp, P.A. (1979) Cell 17, 935-944].

- [0133] Five pairs of synthetic oligonucleotide primers were designed from the C/E, E, E/NS1, NS3, and NS5 domains of the nucleotide sequence of HCV1 to isolate fragments from the J1 and J7 genome. The first set of primers were to isolate the sequence from the core and some of the envelope domain. The second set of primers were to isolate the sequences in the envelope domain. The third set of primers were to isolate a fragment which overlapped the putative envelope and non-structural one, NS1, domains. The fourth and fifth set of primers were used to isolate fragments from non-structural domains three and five, NS3 and NS5. The sequences for the various primers are shown below: 30 The sequence of the primers for the C/E region were:

- 35 21S 5' CGTGCCCCCGCAAGACTGCT 3'  
J80A 5' CCGTCCTCCAGAACCCGGAC 3'

The sequence of the primers for the E region were:

- 40 71S 5' GCCGACCTCATGGGGTACAT 3'  
J132A 5' AACTGCCGACACCACTAAGGC 3'

The sequence of the primers for the E/NS1 region were:

- 45 127S 5' TGGCATGGGATATGATGATG 3'  
166A 5' TTGAACCTGTGGTGATAGAA 3'

The sequence of the primers for the NS3 region were :

- 50 464S 5' GGCTATACCGGCGACTTCGA 3'  
526A 5' GACATGCATGTCATGATGTA 3'

The sequence of the primers for the NS5 region were:

- 55 870S 5' GCTGGAAAGAGGGTCTACTA 3'  
917A 5' GTTCTTACTGCCAGTTGAA 3'

- [0134] 1 µg of the antisense primers, 166A, 526A, or 917A, was added to 10 units of reverse transcriptase (Biorad)

## EP 0 419 182 B1

to synthesize cDNA fragments from the isolated RNA as the template. The cDNA fragments were then amplified by a standard polymerase chain reaction [Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn G.T., Erlich, H.A., and Arnheim, N. (1985) *Science* **230**, 1350-1354] after 1 µg of the appropriate sense primer, 21S, 71S, 127S, 464S or 870S, was added.

5 [0135] The cDNA fragments amplified by the PCR method were gel isolated and cloned by blunt-end ligation into the SmaI site of pUC119 [Vieira, J. and Messing, J. (1987) *Methods in Enzymology* **153**, 3-11] or into the SnaBI site of charomid SB, a derivative of the cloning vector charomid 9-42 [Saito, I. and Stark, G. (1986) *Proc. Natl. Acad. Sci. USA* **83**: 8664-8668]. Clones which contain the fragments of the five viral domains were successfully constructed.

## 10 II

[0136] From the PCR reaction of the Japanese isolates, J1 and J7, three independent clones from each region, C/E, E, E/NS1, NS3, and NS5, have been sequenced by the dideoxy chain termination method.

[0137] Sequence from all regions except C/E has been isolated from the J1 isolate. Sequence from only the C/E 15 region has been isolated from the J7 isolate. Surprisingly, fragments isolated from both isolates are neither longer or shorter than what would be predicted from the HCV1 genome. However, there is heterogeneity between clones containing sequence from the same region. Consequently, a consensus sequence was constructed for each of the domains, C/E, E, E/NS1, NS3 and NS5, as shown respectively in Figures 1 through 5. These differences may be explained as artifacts which occur randomly during the PCR amplification [Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) *Science* **230**, 1350-1354]. Another explanation is that more than 20 one virus genome is present in the plasma of a single healthy carrier and that these genomes are heterogeneous at the nucleotide level.

[0138] To clarify this point, it was determined how many of these nucleotide differences would lead to amino acid changes, using the sequence from the NS3 domain of the J1 isolate as an example. Out of the five nucleotide differences, three fall on the third position of the amino acid codon and do not change the amino acid sequence. Both of the 25 remaining two nucleotide changes fall on the first position of the amino acid codon and generate amino acid changes of threonine to alanine and proline to alanine, all of which are small, neutral amino acid residues. Similarly, when analyzing the nucleotide differences in other domains, many silent and conserved mutations are found. These results suggest that nucleotide sequences of the HCV genomes in the plasma of a single healthy donor are heterogeneous at the 30 nucleotide level.

[0139] In addition, once the consensus sequences for each of the fragments were compiled each sequence was compared to the HCV1 isolate in Figures 6 through 10. In Figure 6 the fragment from the C/E region of the J7 isolate shows a 92.8%, 512/552, nucleotide and 97.4%, 150/154, amino acid homology to the HCV1 isolate. The fragment from the E domain of J1 shows a slightly lower nucleotide and amino acid homology to HCV1 in Figure 7 of 76.2% and 82.9%, 35 respectively. The fragment from the J1 isolate which overlaps the envelope and non-structural one domains shows the lowest homology to HCV1, as seen in Figure 8, where the J1 isolate has a 71.5% nucleotide homology and a 73.5% amino acid homology to HCV1. Figure 9 shows a comparison of the fragment from the NS3 domain of J1 to HCV1. The homology between the nucleotides sequences is 79.8%, while the amino acid homology between the isolates is quite high, 92.2% or 179/194 amino acids. Figure 10 shows the homology between the NS5 sequences from J1 and HCV1. 40 The sequences have a 84.3% nucleotide and 88.7% amino acid homology.

[0140] The vectors described in the examples above were deposited with the Patent Microorganism Depository, Fermentation Institute, Agency of Industrial Science and Technology at 1-3, Higashi 1-chome Tsukuba-chi, Ibaragiken 305, Japan, and will be maintained under the provisions of the Budapest Treaty. The accession numbers and dates of the deposit are listed below, on page 68.

45

## III

[0141] An HCV/J1 clone, J1-1519, was isolated using the essentially the techniques described above. However, the primers used in the isolation were J159S and 199A. The sequences of the oligomeric primers J159S and 199A, which 50 follow, Were based upon those in J1-1216 and in HCV1.

J159S 5' ACT GCC CTG AAC TGC AAT GA 3'  
199A 5' AAT CCA GTT GAG TTC ATC CA 3'

55 [0142] Clone J1-1519 is comprised of an HCV cDNA sequence of 367 nucleotides which spans most of the 5'-half of the NS1 region and which overlaps the E-region clone, J1-1216, by 31 nucleotides. Three independent 10 clones spanning this region were sequenced; the sequences in this region obtained from the three clones were identical. The sequence of the HCV cDNA in J1-1216 (shown in the figure as J1) and the amino acids encoded therein (shown above

the nucleotide sequence) are shown in Figure 13. Figure 13 also shows the sequence differences between J1-1216 in the comparable region of the prototype HCV1 cDNA (indicated in the figure as PT), and the resulting changes in the encoded amino acids. The homology between the J1-1216 and HCV1 cDNA is approximately 70% at the nucleotide level, and about 75% at the amino acid level.

- 5 [0143] A composite of the sequences from the putative core to NS1 region of the J1 isolate is shown in Figure 14; also shown in the figure are the amino acids encoded in the J1 sequence. The variation from the HCV1 prototype sequence is shown in the line below the J1 nucleotide sequence; the dashed lines indicate homologous sequences. The nonhomologous amino acid encoded in the HCV1 prototype sequence is shown below the HCV1 nucleotide sequence.
- 10 [0144] Cloned material containing the J1/1519 HCV cDNA (ps1-1519) has been maintained in DH5 $\alpha$ , and deposited with the Patent Microorganism Depository.

#### IV

- 15 [0145] Several regions of the J1 isolate, including the C200-C100 region from the putative NS3-NS4 region (which encompasses the region encoding the 5-1-1 polypeptide in HSV1 (See EPO Pub. No. 318,216), and the putative NS1 - E region, were amplified using the PCR method. The C200-C100 region includes nucleotides 3799 to 5321 of the prototype HCV1. RNA was extracted as described above, except that extraction was with guanidinium thiocyanate in the presence of Proteinase K and sodium dodecylsulfate (SDS) (Maniatis (1982), supra). The RNA was transcribed into
- 20 HCV cDNA by incubation in a 25  $\mu$ l reaction comprised of 1  $\mu$ M of each primer, 40 units of RNase inhibitor (RNASIN), 5 units of AMV reverse transcriptase, and salts and buffer necessary for the reaction. Amplification of a segment of the HCV cDNA from the designated region was performed utilizing pairs of synthetic oligomer 16-mer primers. PCR amplification was accomplished in three rounds (PCR I, PCR II, and PCRIII). The second and third rounds of PCR amplification (PCR II) utilized different sets of PCR primers; the first PCR reaction was diluted 10-fold and multiple rounds of
- 25 PCR amplification were carried out with the new primers, so that ultimately up to 50% of the products of the first PCR reaction (PCR I) were reamplified. The primers used for the amplification of the regions were the following. These primers, with the exception of J1C200-3 which was derived from the J1 isolate sequence, were derived from the prototype HCV1 sequence.

- 30 Primers for amplification of the "5-1-1" region from NS3-NS4.

#### [0146]

##### PCR I

- 35 511/16A (sense, derived from nucleotides starting at number 1528 of HCV1)

5' AAC AGG CTG CGT GGT C 3'

511/16B (anti-sense, derived from nucleotides ending at 5260 of HCV1)

40

5' AGT TGG TCT GGA CAG C 3'

##### PCR II

- 45 511/35A (sense, the HCV portion derived from nucleotides starting at number 5057 of HCV1, the restriction enzyme site is underlined)

5' CITGAATTC TCG TCT TGT CCG GGA AGC CGG CAA TC 3'

- 50 511/35B (anti-sense, the HCV portion derived from nucleotides ending at number 5233 of HCV1; the restriction enzyme site is underlined)

5' CITGAATTC CCT CTG CCT GAC GGG ACG CGG TCT GC 3'

##### PCRIII

- 55 511/35A (see supra)  
VSNrcZ (antisense, derived from nucleotides ending at number 5804 of HCV1)

5' GTA GTG CGT GGG GGA AAC AT 3'

EP 0 419 182 B1

Primers for amplification of the "NS1/E" region

[0147]

5 PCR I  
J1(E2)3 (sense, the HCV portion derived from nucleotides starting at number 953 of HCV1, the restriction enzyme site is underlined)  
5' CTTAGAATTC TGG CAT GGG ATA TGA TGA TG 3'  
10 J1(E)4 (sense, the HCV portion derived from nucleotides starting at number 1087 of HCV1, the restriction enzyme site is underlined)  
5' CTTAGAATTC TCC ATG GTG GGG AAC TGG GC 3'  
15 J1rc12 (anti-sense, the HCV portion derived from nucleotides ending at 1995 of HCV1, the restriction enzyme site is underlined)  
5' CTTGAATTC TAA CGG GCT GAG CTC GGA 3'  
20 J1rc13 (anti-sense, the HCV portion derived from nucleotides ending at 1941 of HCV1, the restriction enzyme site is underlined)  
5' CTTAGAATTC CGT CCA GTT GCA GGC AGC TTC 3'  
25 PCR II  
J1rc13 (see supra)  
J1I2-1 (sense, the HCV portion is derived from nucleotides starting at number 1641 of HCV1, the restriction enzyme site is underlined)  
30 5' CTTGAATTC CAA CTG GTT CGG CTG TAC A 3'  
J1I2-2 (sense, the HCV portion is derived from nucleotides starting at number 1596 of HCV1, the restriction enzyme site is underlined)  
35 5' TGA GAC GGA CGT GCT GCT CCT 3'

Primers for the C200-C100 region of the "NS3-NS4" region

40 [0148]

PCR I  
J1C200-1 (sense, derived from nucleotides starting at number 3478 of HCV1)  
45 5' TCC TAC TTG AAA GGC TC 3'  
J1C200-3 (anti-sense, derived from nucleotides ending at number 4402 of HCV1)  
5' GGA TCC AAG CTG AAA TCG AC 3'  
50 J1rc52 (anti-sense, the HCV portion derived from nucleotides ending at 5853 of HCV1, the restriction enzyme site is underlined)  
5' CTTAGAATTC GAG GCT GCT GAG ATA GGC AGT 3'  
55 511/16A (see above).  
PCR II  
J1C200-2 (sense, the HCV portion derived from nucleotides starting at number 3557 of HCV1, the restriction

enzyme site is underlined)

5' CTTGAATTC CCC GTG GAG TGG CTA AGG CGG TGG ACT 3'

5 J1C200-4 (anti-sense, the HCV portion derived from nucleotides ending at 4346 of HCV1, the restriction enzyme site is underlined)

5' CTTGAATTC TCG AAG TCG CCG GTA TAG CCG GTC ATG 3'

10 511/35A (see above)

J1rc51 (anti-sense, the HCV portion derived from nucleotides ending at 5826 of HCV1, the restriction enzyme site is underlined)

15 5' CTTAGAATTC GGC AGC TGC ATC GCT CTC CGG CAC 3'

The amplified HCV cDNAs were either sequenced directly without cloning, and/or were cloned. Sequencing was accomplished using an asymmetric PCR technique, essentially as described in Shyamala and Ames, J. Bacteriology 171:1602 (1989). In this technique, amplification of the cDNA is carried out with a limiting concentration of one of the primers (usually in a ratio of about 1:50) in order to get preferential amplification of one strand. The preferentially amplified strand is then sequenced by the dideoxy chain termination method.

The primers used for asymmetric sequencing by the PCR method were the following. For the NS1 region: J1IZ-1 and J1rc13 (sequenced with both); J1IZ-2, J1rc13 (confirmed on both strands). For the NS3-NS4 region, which includes the C200-C100 N-terminal region, C200-C100 C-terminal region, and the 5-1-1 region: J1C200-2 and J1C200-7 (for the N-terminal region of C200-C100), and J1C200-4 and J1C200-6 (for the C200-C100 C-terminal region); and 511/35A and hep 4 (for the 5-1-1 region). The sequences for J1C200-2, J1C200-4, and 511/35A are shown supra; the sequences of hep 4, J1C200-6, and J1C200-7 are the following.

hep 4 (derived from nucleotides starting at number 5415 of HCV1)

30 5' TT GGC TAG TGG TTA GTG GGC TGG TGA CAG 3'

J1C200-6 (the HCV portion derived from nucleotides starting at number 3875 of HCV1, the restriction enzyme site is underlined)

35 5' CTTGAATTC CGT ACT CCA CCT ACG GCA AGT TCC TT 3'

J1C200-7 (the HCV portion derived from nucleotides starting at number 3946 of HCV1, the restriction enzyme site is underlined)

40 5' CTTGAATTC GTG GCA TCC GTG GAG TGG CAC TCG TC 3'

[0149] The sequences obtained by asymmetric sequencing of the "NS1" region, the C200-C100 region, and the 5-1-1 region are shown in Figure 15, and Figure 16, respectively. In the figures, the amino acids encoded in the J1 sequence are shown above the J1 nucleotide sequence. The differences between the J1 sequence and the HCV1 prototype nucleotide sequence is shown below the J1 sequence (the dashes indicate homologous nucleotides in both sequences). The encoded amino acids which differ in the HCV1 prototype sequence are shown below the HCV1 nucleotide sequence.

[0150] HCV cDNAs from the NS1 region, the C200-C100 region, and the 5-1-1 region were cloned. A 300 bp and a 230 bp fragment from the putative NS1 region, were cloned into a derivative of the commercially available vector, pGEM-3Z, in host HB101, and deposited with the ATCC as AW-300bp. The derivative vectors maintain the original pGEM-3Z polylinkers, an intact Amp<sup>r</sup> gene, and the genes required for replication in *E. coli*. The HCV cDNA fragments may be removed with SacI and XbaI. HCV cDNAs containing 770 bp N-terminal fragments of C200 were cloned into pM1E in HB101, 12 clones were pooled and deposited with the ATCC as AW-770bp-N; the HCV cDNA may be removed from the vector with HaeII. The resultant HaeII fragment will contain vector DNA of 300 bp and 250 bp at the 5' and 3' ends, respectively. HCV cDNAs containing 700 bp C-terminal fragments of C200 (AW-700bp-C) were cloned into M13mp10 and maintained in host DH5 $\alpha$ -F<sup>-</sup>; cloning was into the vector polylinker site. The resultant phage were pooled, and deposited with the ATCC on September 11, 1990 as AW-700bp-N or AW-700bp-C. HCV cDNA from J1 equivalent to the 5-1-1 region of HCV1 was cloned into mp19EcoR1 site, and maintained in DH5 $\alpha$ -F<sup>-</sup>. Several m13

# EP 0 419 182 B1

phage superanants from this cloning were pooled and deposited with the ATCC as J1 5-1-1, on September 11, 1990. The HCV cDNAs may be obtained from the phage by treatment with EcoRI. Accession numbers for J1 5-1-1 and AW-700bp-N or AW-700bp-C may be obtained by telephoning the ATCC at (301) 881-2600.

[0151] The above-described cloned material was deposited with the American Type Culture Collection (ATCC).

## V

[0152] An HCV cDNA library containing sequences of the putative "NS1" region of the J1 isolate was created by directional cloning in  $\lambda$ -gt22. The "NS1" region extends from about nucleotide 1460 to about nucleotide 2730 using the numbering system of the HCV1 prototype nucleic acid sequence, where nucleotide 1 is the first nucleotide of the initiating methionine codon for the putative polyprotein. The cloning was accomplished using essentially the method described by Han and Rutter in GENETIC ENGINEERING, Vol 10 (J.K. Setlow, Ed., Plenum Publishing Co., 1988), except that the primers for the synthesis of the first and second strand of HCV cDNA were JHC67 and JHC68, respectively, and the source of RNA was the J1 plasma. In this method the RNA is extracted with guanidium thiocyanate at a low temperature. The RNA is then converted to full length cDNA, which is cloned in a defined orientation relative to the lacZ promoter in  $\lambda$ -phage. Using this method, the HCV cDNAs to J1 RNA were inserted into the NotI site of  $\lambda$ -gt22. The presence of "NS1" sequences in the library was detected using as probe, Abx54.

[0153] The sequence of a region of "NS1" downstream from the region shown in Figure 14, but which overlaps the region by about 20 nucleotides, was determined using the asymmetric sequencing technique described above, but substituting as primers for PCR amplification, Alx 61 and Alx 62. The resulting sequence is shown in Figure 17. (It should be noted that the PCR amplification was of a region from about nucleotide 1930 to about nucleotide 2340; this region is also encompassed in the sequence shown in Figure 15). The sequences of the primers and probes used to obtain the HCV cDNA library in  $\lambda$ -gt22, and to sequence the portion of the "NS1" region were the following.

### JHC 67

5' GACGC GGCCG CCTCC GTGTC CAGCG CGT 3'

### JHC 68

5' CGTGC GGCCG CAAGA CTGCT AGCCG AGGT 3'

### ALX 61

5' ACCTG CCACT GTGTA GTGGT CAGCA GTAAC 3'

### ALX 62

5' ACGGA CGTCT TCGTC CTTAACAATA CCAGG 3'

### ALX 54

5' GAACT TTGCG ATCTG GAAGACAGGG ACAGG 3'

[0154] A 400 bp fragment of J1 HCV cDNA derived from the sequenced region was cloned into pGEM3z and maintained in HB101; the HCV cDNA may be removed from the vector with SacI and XbaI. Host cells transformed with the vector (JH-400bp) have been deposited with the ATCC.

[0155] A pooled cDNA library was created from the J1 serum; the pooled library spans the J1 genome and is identified as HCV-J1  $\lambda$  gt22. The pooled cDNA library was created by pooling aliquots of 11 individual cDNA libraries, which had been prepared using the directional cloning technique described above, except that the libraries were created from primers which were designed to yield HCV cDNAs which spanned the genome. The primers were derived from the sequence of HCV1, and included JHC 67 and JHC 68. The HCV cDNAs were inserted into the NotI site of  $\lambda$ -gt22. The pooled cDNA library, HCV-J1  $\lambda$  gt22, has been deposited with the ATCC.

## VI

[0156] The sequence of a region of the polynucleotide upstream of that shown in Figure 14 was determined. This region begins at nucleotide -267 with respect to the HCV-1 (See Figure 12) and extends for 560 nucleotides. Sequenc-

ing was accomplished by preparing HCV cDNA from RNA extracted from J1 serum, and amplifying the HCV cDNA using the PCR method.

[0157] RNA was extracted from 100  $\mu$ l of serum following treatment with proteinase K and sodium dodecylsulfate (SDS). The samples were extracted with phenol-chloroform, and the RNA precipitated with ethanol.

5 [0158] HCV cDNA from the J1 isolate was prepared by denaturing the precipitated RNA with 0.01M MeHgOH; after ten minutes at room temperature, 2-mercaptoethanol was added to sequester the mercury ions. Immediately, the mix for the first strand of cDNA synthesis was added, and incubation was continued for 1 hr at 37°C. The conditions for the synthesis of the anti-sense strand were the following: 50 mM Tris HCl, pH 8.3, 75 mM KCl, 3 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 500 $\mu$ M each deoxynucleotide triphosphate, 250 pmol specific antisense cDNA primer r25, 250 units MMLV  
10 reverse transcriptase. In order to synthesize the second strand (sense), the synthesis reaction components were added, and incubated for one hour at 14°C. The components for the second strand reaction were as follows: 14 mM Tris HCl, pH 8.3, 68 mM KCl, 7.5 mM ammonium sulfate, 3.5 mM MgCl<sub>2</sub>, 2.8 mM dithiothreitol, 25 units DNA polymerase I, and one unit RNase H. The reactions were terminated by heating the samples to 95°C for 10 minutes, followed by cooling on ice.

15 [0159] The HCV cDNA was amplified by two rounds of PCR. The first round was accomplished using 20  $\mu$ l of the cDNA mix. The conditions for the PCR reaction were as follows: 10 mM Tris HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.002% gelatin, 200 mM each of the deoxynucleotide triphosphates, and 2.5 units Amplitaq. The PCR thermal cycle was as follows: 94°C one minute, 50°C one minute, 72°C one minute, repeated 40 times followed by seven minutes at 72°C. The second round of PCR was accomplished using nested primers (i.e. primers which bound to an internal region  
20 of the first round of PCR amplified product) to increase the specificity of the PCR products. One percent of the first PCR reaction was amplified essentially as the first round, except that the primers were substituted, and the second step in the PCR reaction was at 60°C instead of 50°C. The primers used for the first round of PCR were ALX90 and r14. The primers used for the second round of PCR were r14 and p14.

[0160] The sequences of the primers for the synthesis of HCV cDNA and for the PCR method were the following.

25 r25

5' ACC TTA CCC AAA TTG CGC GAC CTA 3'

30 ALX90

5' CCA TGA ATC ACT CCC CTG TGA GGA ACT A 3'

35 r14

5' GGG CCC CCAG CTA GGC CGA GA 3'

p14

40 5' AAC TAC TGT CTT CAC GCA GAA AGC 3'

[0161] The PCR products were gel purified, the material which migrated as having about 615 bp was isolated, and sequenced by a modification of the Sanger dideoxy chain termination method, using <sup>32</sup>P-ATP as label. In the modified  
45 method, the sequence replication was primed using P32 and R31 as primers; the double stranded DNA was melted for 3 minutes at 95°C prior to replication, and the synthesis of labeled dideoxy terminated polynucleotides was catalyzed by Bst polymerase (obtained from BioRad Corp.), according to the manufacturer's directions. The sequencing was performed using 500ng to 1  $\mu$ g of PCR product per sequencing reaction.

[0162] The primers P32 (sense) and R31 (antisense) were derived from nucleotides -137 to -115 and from nucleotides 192 to 173, respectively, of the HCV1 sequence. The sequences of the primers are the following.

50 P32 primer

5' AAC CCG CTC AAT GCC TGG AGA TT 3'

55 R31 primer

5' GGC CGX CGA GCC TTG GGG AT 3'

EP 0 419 182 B1

where X = A or G

[0163] The sequence of the region in the J1 isolate which encompasses the 5'-untranslated region as well as a part of the region of the putative "Core" is shown in Figure 18. In the figure, amino acids encoded in the J1 sequence are shown above the nucleotide sequence. The sequence of the prototype HCV1 is shown below the J1 sequence; the dashes indicate sequence homology with J1. The differing amino acids encoded in the HCV1 sequence are shown below the HCV1 sequence.

[0164] An HCV cDNA fragment which is a representative of the 600 bp J1 sequence described above (TC 600bp) was cloned into pGEM3Z and maintained in host HB101; the HCV cDNA fragment may be removed with SacI and XbaI. This material is on deposit with the ATCC.

[0165] Patent Microorganism Depository-deposited under Budapest Treaty terms.

15

20

25

30

Deposited Materials	Accession Number	Deposit Date
<u>E. coli</u> DH5/pS1-8791a (This clone contains 427 bp of the HS5 domain of J1)	BP-2593	9/15/1989
<u>E. coli</u> HB101/pU1-1216c (This clone contains 351 bp of the E/NS1 domains of J1)	BP-2594	9/15/1989
<u>E. coli</u> HB101/pU1-4652d (This clone contains 583 bp of the NS3 domain of J1)	BP-2595	9/15/1989
<u>E. coli</u> DH5α/pS1-713c (This clone contains 580 bp of the E domain of J1)	BP-2637	11/1/1989
<u>E. coli</u> DH5α/pS7-28c (This clone contains 552 bp of the C/E domain of J7)	BP-2638	11/1/1989
<u>E. coli</u> DH5α/ps1-1519	BP3081	8/30/90

[0166] The following vectors described in the Examples were deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Dr., Rockville, Maryland 20852, and have been assigned the following Accession Numbers. The deposits were made under the terms of the Budapest Treaty.

40

45

50

Deposited Materials	Accession Number	Deposit Date
TC-600BP (in <u>E. coli</u> HB101/pGEM3Z)	68393	9/11/90
JH-400bp (in <u>E. coli</u> HB101/pGEM3Z)	68394	9/11/90
AW-300bp (in <u>E. coli</u> HB101/pGEM3Z)	68392	9/11/90
AW-770bp-N (in <u>E. coli</u> HB101/pM1E)	68395	9/11/90
AW-700bp-C or AW-700bp-N (in <u>E. coli</u> DH5α-F/M13mp10)		
J1 5-1-1 (in <u>E. coli</u> DH5α-F/M13mp10)		
HCV-J1 λ gt22	40884	9/6/90

[0167] These deposits are provided for the convenience of those skilled in the art. These deposits are neither an admission that such deposits are required to practice the present invention nor that equivalent embodiments are not within the skill of the art in view of the present disclosure. The public availability of these deposits is not a grant of a license to make, use or sell the deposited materials under this or any other patent. The nucleic acid sequences of the deposited materials are incorporated in to present disclosure by reference, and are controlling if in conflict with any sequences described herein.

[0168] While the present invention has been described by way specific examples for the benefit of those in the field,



the scope of the invention is not limited as additional embodiments will be apparent to those of skill in the art from the present disclosure.

# Claims

5

Claims for the following Contracting States : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, LU, NL, SE

10

1. A polynucleotide in substantially isolated form comprising a nucleotide sequence of at least 15 nucleotides from a J-1 HCV isolate, said J-1 HCV isolate having at least 90% nucleotide sequence homology with the J-1 sequence of any one of Figures 7 to 10 or 13 to 18, wherein said nucleotide sequence of at least 15 nucleotides is distinct from the nucleotide sequence of HCV isolate HCV-1

15

2. A polynucleotide according to claim 1 wherein the J-1 HCV isolate has at least 95% homology with the J-1 HCV sequence of any one of Figures 7 to 10 or 13 to 18.

20

3. A polynucleotide according to claim 1 wherein the J-1 HCV isolate has 100% homology with the J-1 HCV sequence of any one of Figures 7 to 10 or 13 to 18.

4. A polynucleotide according to any one of the preceding claims which comprises at least 20 nucleotides.

25

5. A method of detecting HCV polynucleotides in a test sample comprising:

- (a) providing a polynucleotide as defined in any one of claims 1 to 4 as a probe;
- (b) contacting the test sample and the probe under conditions that allow for the formation of a polynucleotide duplex between the probe and its complement in the absence of substantial polynucleotide duplex formation between the probe and non-HCV polynucleotide sequences present in test sample; and
- (c) detecting any polynucleotide duplexes comprising the probe.

30

6. A polynucleotide comprising a sequence of at least 15 nucleotides from a J-1 HCV isolate present in any one of plasmids pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW-770-BP-N, or J1 5-1-1 deposited under accession numbers BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 and 40884 respectively, wherein said nucleotide sequence of at least 15 nucleotides is distinct from the nucleotide sequence of HCV isolate HCV-1.

35

7. A purified polypeptide comprising an amino acid sequence which:

- (a) is encoded by a nucleotide sequence as defined in any one of claims 1 to 4 or in the HCV sequences deposited and defined in claim 6, said coding being in frame with the corresponding amino acid sequences set out in Figures 7 to 10 or 13 to 18;
- (b) comprises an antigenic determinant; and
- (c) is distinct from the sequence of the polypeptides encoded by the HCV isolate HCV-1

40

8. A polypeptide according to claim 7 which comprises at least 10 amino acids.

45

9. A polypeptide according to claim 7 which comprises at least 15 amino acids.

10. A polypeptide according to any one of claims 7 to 9 immobilised on a solid support.

50

11. An immunoassay for detecting the presence of anti-HCV antibodies in a test sample which comprises:

- (a) incubating the test sample under conditions that allow the formation of an antigen-antibody complex to be formed with a polypeptide as defined in any one of claims 7 to 10, wherein the polypeptide is not immunologically cross-reactive with HCV-1; and
- (b) detecting any antigen antibody complexes formed.

55

12. An immunoassay according to claim 11 wherein the test sample comprises human blood or a fraction thereof.

Claims for the following Contracting State : ES

1. A method of preparing a polynucleotide in substantially isolated form which comprises a nucleotide sequence of at least 15 nucleotides from a J-1 HCV isolate, said J-1 HCV isolate having at least 90% nucleotide sequence homology with the J-1 sequence of any one of Figures 7 to 10 or 13 to 18, wherein said nucleotide sequence of at least 15 nucleotides is distinct from the nucleotide sequence of HCV isolate HCV-1, said method comprising:
  - a) chemical synthesis by methods known *per se*; or
  - b) DNA replication by methods known *per se*; or
  - c) transcription or reverse transcription by methods known *per se*; or
  - d) restriction enzyme digestion and ligation into a vector by recombinant DNA techniques known *per se*.
2. A method of preparing a polypeptide comprising an amino acid sequence which is encoded by a nucleotide sequence as defined in claim 1, said method comprising:
  - (i) chemical synthesis by methods known *per se*; or
  - (ii) translation from the corresponding nucleic acid sequence by methods known *per se*; or
  - (iii) expression of a recombinant expression system comprising the nucleotide sequence encoding the polypeptide by methods known *per se*; or
  - (iv) isolation from virus by methods known *per se*.
3. A method of detecting HCV polynucleotides in a test sample comprising:
  - (a) providing a polynucleotide as defined in claim 1 as a probe;
  - (b) contacting the test sample and the probe under conditions that allow for the formation of a polynucleotide duplex between the probe and its complement in the absence of substantial polynucleotide duplex formation between the probe and non-HCV polynucleotide sequences present in test sample; and
  - (c) detecting any polynucleotide duplexes comprising the probe.
4. An immunoassay for detecting the presence of anti-HCV antibodies in a test sample which comprises:
  - (a) incubating the test sample under conditions that allow the formation of an antigen-antibody complex to be formed with a polypeptide which is encoded by a nucleotide sequence as defined in claim 1 or a sequence of at least 15 nucleotides from a J-1 HCV isolate present in any one of plasmids pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW-770-BP-N or J1 5-1-1 deposited under accession numbers BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 and 40884 respectively, wherein said nucleotide sequence of at least 15 nucleotides is distinct from the nucleotide sequence of HCV isolate HCV-1, said coding being in frame with the corresponding amino acid sequences set out in Figures 7 to 10 or 13 to 18, which polypeptide comprises an antigenic determinant and is distinct from the sequence of the polypeptide encoded by the HCV isolate HCV-1 and is not immunologically cross-reactive with HCV-1; and
  - (b) detecting any antigen antibody complexes formed.
5. An immunoassay according to claim 6 wherein the test sample comprises human blood or a fraction thereof.
6. A polynucleotide in substantially isolated form comprising a nucleotide sequence of at least 15 nucleotides from a J-1 HCV isolate, said J-1 HCV isolate having at least 90% nucleotide sequence homology with the J-1 sequence of any one of Figures 7 to 10 or 13 to 18, wherein said nucleotide sequence of at least 15 nucleotides is distinct from the nucleotide sequence of HCV isolate HCV-1.
7. A polynucleotide according to claim 6 wherein the J-1 HCV isolate has at least 95% homology with the J-1 HCV sequence of any one of Figures 7 to 10 or 13 to 18.
8. A polynucleotide according to claim 6 wherein the J-1 HCV isolate has 100% homology with the J-1 HCV sequence of any one of Figures 7 to 10 or 13 to 18.
9. A polynucleotide according to any one of claims 6 to 8 which comprises at least 20 nucleotides.

10. A polynucleotide comprising a sequence of at least 15 nucleotides from a J-1 HCV isolate present in any one of plasmids pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW-770-BP-N or J1 5-1-1 deposited under accession numbers BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 and 40884 respectively, wherein said nucleotide sequence of at least 15 nucleotides is distinct from the nucleotide sequence of HCV isolate HCV-1.
11. A purified polypeptide comprising an amino acid sequence which:
- (a) is encoded by a nucleotide sequence as defined in any one of claims 6 to 9 or in the HCV sequences deposited and defined in claim 10, said coding being in frame with the corresponding amino acid sequences set out in Figures 7 to 10 or 13 to 18;
  - (b) comprises an antigenic determinant; and
  - (c) is distinct from the sequence of the polypeptide encoded by the HCV isolate HCV-1.
12. A polypeptide according to claim 11 which comprises at least 10 amino acids.
13. A polypeptide according to claim 11 which comprises at least 15 amino acids.
14. A polypeptide according to any one of claims 11 to 13 immobilised on a solid support.

#### Patentansprüche

Patentansprüche für folgende Vertragsstaaten : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, LU, NL, SE

1. Polynucleotid in im wesentlichen isolierter Form, umfassend eine Nucleotidsequenz aus mindestens 15 Nucleotiden aus einem J-1 HCV-Isolat, wobei das J-1 HCV-Isolat eine mindestens 90%ige Nucleotidsequenzhomologie mit der J-1-Sequenz einer der Figuren 7 bis 10 oder 13 bis 18 besitzt, wobei sich die Nucleotidsequenz aus mindestens 15 Nucleotiden von der Nucleotidsequenz des HCV-Isolats HCV-1, wie in Figur 12 gezeigt, unterscheidet.
2. Polynucleotid nach Anspruch 1, wobei das J-1 HCV-Isolat eine mindestens 95%ige Homologie mit der J-1 HCV-Sequenz einer der Figuren 7 bis 10 oder 13 bis 18 besitzt.
3. Polynucleotid nach Anspruch 1, wobei das J-1 HCV-Isolat eine 100%ige Homologie mit der J-1 HCV-Sequenz einer der Figuren 7 bis 10 oder 13 bis 18 besitzt.
4. Polynucleotid nach einem der vorstehende Ansprüche, das mindestens 20 Nucleotide umfaßt.
5. Verfahren zum Nachweis von HCV-Polynucleotiden in einer Testprobe, umfassend:
  - (a) Bereitstellen eines Polynucleotids, wie in einem der Ansprüche 1 bis 4 definiert, als Sonde,
  - (b) In-Kontakt-Bringen der Testprobe und der Sonde unter Bedingungen, die die Bildung von Polynucleotid-Doppelsträngen zwischen der Sonde und deren Komplementärstrang in Abwesenheit einer wesentlichen Polynucleotid-Doppelstrangbildung zwischen der Sonde und Nicht-HCV-Polynucleotidsequenzen, die in der Testprobe vorhanden sind, erlauben; und
  - (c) Nachweis von allen Polynucleotid-Doppelsträngen, die die Sonde umfassen.
6. Polynucleotid, umfassend eine Sequenz aus mindestens 15 Nucleotiden aus einem J-1 HCV-Isolat, das in einem der Plasmide pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW-770-BP-N oder J1 5-1-1 hinterlegt unter den Hinterlegungsnummern BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 bzw. 40884 vorhanden ist, wobei sich die Nucleotidsequenz aus mindestens 15 Nucleotiden von der Nucleotidsequenz des HCV-Isolats HCV-1 unterscheidet.
7. Gereinigtes Polypeptid, umfassend eine Aminosäuresequenz, die
  - (a) von einer Nucleotidsequenz, die in einem der Ansprüche 1 bis 4 definiert ist, oder von den hinterlegten und in Anspruch 6 definierten HCV-Sequenzen codiert wird, wobei die Codierung im Leseraster mit den entspre-

chenden Aminosäuresequenzen, die in den Figuren 7 bis 10 oder 13 bis 18 dargestellt sind, ist;

(b) eine antigene Determinante umfaßt; und

5 (c) sich von der Sequenz der von dem HCV-Isolat HCV-1 codierten Polypeptide unterscheidet.

8. Polypeptid nach Anspruch 7, das mindestens 10 Aminosäuren umfaßt.

9. Polypeptid nach Anspruch 7, das mindestens 15 Aminosäuren umfaßt.

10

10. Polypeptid nach einem der Ansprüche 7 bis 9, das auf einem festen Träger immobilisiert ist.

11. Immunoassay zum Nachweis der Anwesenheit von Anti-HCV-Antikörpern in einer Testprobe, wobei man

15

(a) die Testprobe unter Bedingungen inkubiert, die die Bildung eines zu bildenden Antigen-Antikörper-Komplexes mit einem Polypeptid, wie in einem der Ansprüche 7 bis 10 definiert, erlauben, wobei das Polypeptid mit HCV-1 nicht immunologisch kreuzreaktiv ist; und

(b) alle gebildeten Antigen-Antikörper-Komplexe nachweist.

20

12. Immunoassay nach Anspruch 11, worin die Testprobe menschliches Blut oder eine Fraktion davon umfaßt.

#### Patentansprüche für folgenden Vertragsstaat : ES

25

1. Verfahren zur Herstellung eines Polynucleotids in im wesentlichen isolierter Form, das eine Nucleotidsequenz aus mindestens 15 Nucleotiden aus einem J-1 HCV-Isolat umfaßt, wobei das J-1 HCV-Isolat eine mindestens 90%ige Nucleotidsequenzhomologie mit der J-1-Sequenz einer der Figuren 7 bis 10 oder 13 bis 18 besitzt, wobei sich die Nucleotidsequenz aus mindestens 15 Nucleotiden von der Nucleotidsequenz des HCV-Isolats HCV-1, unterscheidet, wobei das Verfahren umfaßt:

30

a) chemische Synthese nach an sich bekannten Verfahren; oder

b) DNA-Replikation nach an sich bekannten Verfahren; oder

35

c) Transkription oder Umkehrtranskription nach an sich bekannten Verfahren; oder

d) Restriktionsenzymspaltung und Ligierung in einem Vektor durch an sich bekannte rekombinante DNA-Techniken.

40

2. Verfahren zur Herstellung eines Polypeptids, umfassend eine Aminosäuresequenz, die von einer Nucleotidsequenz, wie in Anspruch 1 definiert, codiert wird, wobei das Verfahren umfaßt:

(i) chemische Synthese nach an sich bekannten Verfahren; oder

45

(ii) Translation von der entsprechenden Nucleinsäuresequenz nach an sich bekannten Verfahren; oder

(iii) Expression eines rekombinanten Expressionssystems, das die Nucleotidsequenz umfaßt, die das Polypeptid codiert, nach an sich bekannten Verfahren; oder

50

(iv) Isolation aus einem Virus nach an sich bekannten Verfahren.

3. Verfahren zum Nachweise von HCV-Polynucleotiden in einer Testprobe, umfassend:

55

(a) Bereitstellen eines Polynucleotids, wie in Anspruch 1 definiert, als Sonde,

(b) In-Kontakt-Bringen der Testprobe und der Sonde unter Bedingungen, die die Bildung eines Polynucleotid-Doppelstranges zwischen der Sonde und seinem Komplementärstrang in Abwesenheit einer wesentlichen Bildung von Polynucleotid-Doppelsträngen zwischen der Sonde und den in der Testprobe vorhandenen Nicht-

HCV-Polynucleotidsequenzen erlauben; und

(c) Nachweis aller Polynucleotid-Doppelstränge, die die Sonde umfassen.

5 4. Immunoassay zum Nachweis der Anwesenheit von Anti-HCV-Antikörpern in einer Testprobe, umfassend:

(a) Inkubieren der Testprobe unter Bedingungen, die die Bildung eines zu bildenden Antigen-Antikörper-Komplexes mit einem Polypeptid erlauben, das von einer Nucleotidsequenz, wie in Anspruch 1 definiert, oder einer Sequenz aus mindestens 15 Nucleotiden aus einem J-1 HCV-Isolat, das in einem der Plasmide pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW-770-BP-N oder J1 5-1-1, 10 die unter den Hinterlegungsnummern BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 bzw. 40884 hinterlegt wurden, vorhanden ist, codiert wird, wobei sich die Nucleotidsequenz aus mindestens 15 Nucleotiden von der Nucleotidsequenz des HCV-Isolats HCV-1 unterscheidet, wobei die Codierung im Leseraster mit den entsprechenden Aminosäuresequenzen der Figuren 7 bis 10 oder 13 bis 18 ist, wobei 15 das Polypeptid eine antigene Determinante umfaßt und sich von der Sequenz des von dem HCV-Isolat HCV-1 codierten Polypeptids unterscheidet und mit HCV-1 nicht immunologisch kreuzreaktiv ist; und

(b) Nachweis aller gebildeten Antigen-Antikörper-Komplexe.

20 5. Immunoassay nach Anspruch 6, worin die Testprobe menschliches Blut oder eine Fraktion davon umfaßt.

6. Polynucleotid in im wesentlichen isolierter Form, umfassend eine Nucleotidsequenz von mindestens 15 Nucleotiden aus einem J-1 HCV-Isolat, wobei das J-1 HCV-Isolat mindestens 90% Nucleotidsequenz-Homologie mit der J-1-Sequenz einer der Figuren 7 bis 10 oder 13 bis 18 besitzt, wobei sich die Nucleotidsequenz von mindestens 15 25 Nucleotiden von der Nucleotidsequenz des HCV-Isolats HCV-1 unterscheidet.

7. Polynucleotid nach Anspruch 6, wobei das J-1 HCV-Isolat eine mindestens 95%ige Homologie mit der J-1 HCV-Sequenz einer der Figuren 7 bis 10 oder 13 bis 18 besitzt.

30 8. Polynucleotid nach Anspruch 6, wobei das J-1 HCV-Isolat eine 100%ige Homologie mit der J-1 HCV-Sequenz einer der Figuren 7 bis 10 oder 13 bis 18 besitzt.

9. Polynucleotid nach einem der Ansprüche 6 bis 8, das mindestens 20 Nucleotide umfaßt.

35 10. Polynucleotid, umfassend eine Sequenz aus mindestens 15 Nucleotiden aus einem J-1 HCV-Isolat, das in einem der Plasmide pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW-770-BP-N oder J1 5-1-1, hinterlegt unter den Hinterlegungsnummern BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 bzw. 40884, vorhanden ist, wobei sich die Nucleotidsequenz aus mindestens 15 Nucleotiden von der Nucleotidsequenz des HCV-Isolats HCV-1 unterscheidet. 40

11. Gereinigtes Polypeptid, umfassend eine Aminosäuresequenz, die

(a) von einer Nucleotidsequenz, wie in einem der Ansprüche 6 bis 9 definiert oder den HCV-Sequenzen, die hinterlegt wurden und in Anspruch 10 definiert sind, codiert wird, wobei die Codierung im Leseraster mit den entsprechenden Aminosäuresequenzen, die in den Figuren 7 bis 10 oder 13 bis 18 dargestellt sind, erfolgt; 45

(b) eine antigene Determinante umfaßt; und

(c) sich von der Sequenz des von dem HCV-Isolat HCV-1 codierten Polypeptids unterscheidet. 50

12. Polypeptid nach Anspruch 11, das mindestens 10 Aminosäuren umfaßt.

13. Polypeptid nach Anspruch 11, das mindestens 15 Aminosäuren umfaßt.

55 14. Polypeptid nach einem der Ansprüche 11 bis 13, das auf einem festen Träger immobilisiert ist.

**Revendications**

Revendications pour les Etats contractants suivants : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, LU, NL, SE

- 5 1. Polynucléotide sous une forme sensiblement isolée, comprenant une séquence nucléotidique d'au moins 15 nucléotides d'un isolat HCV J-1, ledit isolat HCV J-1 ayant une homologie de séquence nucléotidique d'au moins 90 % avec la séquence J-1 de l'une quelconque des figures 7 à 10 ou 13 à 18, dans lequel ladite séquence nucléotidique d'au moins 15 nucléotides est différente de la séquence nucléotidique de l'isolat HCV-1 de HCV, telle que représentée à la figure 12.
- 10 2. Polynucléotide selon la revendication 1 dans lequel l'isolat HCV J-1 a une homologie d'au moins 95 % avec la séquence de HCV J-1 de l'une quelconque des figures 7 à 10 ou 13 à 18.
- 15 3. Polynucléotide selon la revendication 1 dans lequel l'isolat HCV J-1 a une homologie de 100 % avec la séquence de HCV J-1 de l'une quelconque des figures 7 à 10 ou 13 à 18.
4. Polynucléotide selon l'une quelconque des revendications précédentes qui comprend au moins 20 nucléotides.
- 20 5. Méthode de détection de polynucléotides de HCV dans un échantillon de test comprenant :
  - (a) la fourniture d'un polynucléotide tel que défini dans l'une quelconque des revendications 1 à 4 comme sonde ;
  - (b) la mise en contact de l'échantillon de test et de la sonde dans des conditions qui permettent la formation d'un duplex polynucléotidique entre la sonde et son complément en l'absence de formation appréciable de duplex polynucléotidiques entre la sonde et des séquences polynucléotidiques autres que de HCV présentes dans l'échantillon de test ; et
  - (c) la détection de tous duplex polynucléotidiques comprenant la sonde.
- 25 6. Polynucléotide comprenant une séquence d'au moins 15 nucléotides d'un isolat HCV J-1 présent dans l'un quelconque des plasmides pS1-8791a, bU1-1216c, bu1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW 770-BP-N ou J1 5-1-1 déposés sous les numéros d'accès BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 et 40884 respectivement, ladite séquence nucléotidique d'au moins 15 nucléotides étant différente de la séquence nucléotidique de l'isolat HCV-1 de HCV.
- 30 7. Polypeptide purifié comprenant une séquence d'acides aminés qui :
  - (a) est codée par une séquence nucléotidique telle que définie dans l'une quelconque des revendications 1 à 4 ou dans les séquences de HCV déposées et définies à la revendication 6, ladite séquence codante étant en phase avec les séquences d'acides aminés correspondantes présentées dans les figures 7 à 10 ou 13 à 18 ;
  - (b) comprend un déterminant antigénique ; et
  - (c) est distincte de la séquence des polypeptides codés par l'isolat HCV-1 de HCV.
- 35 8. Polypeptide selon la revendication 7 qui comprend au moins 10 acides aminés.
- 40 9. Polypeptide selon la revendication 7 qui comprend au moins 15 acides aminés.
10. Polypeptide selon l'une quelconque des revendications 7 à 9 immobilisé sur un support solide.
- 45 11. Test immunologique destiné à détecter la présence d'anticorps anti-HCV dans un échantillon de test qui comprend :
  - (a) l'incubation de l'échantillon de test dans des conditions qui permettent la formation d'un complexe antigène-anticorps à former avec un polypeptide tel que défini dans l'une quelconque des revendications 7 à 10, le polypeptide ne produisant pas de réaction immunologique croisée avec HCV-1 ; et
  - (b) la détection de tout complexe antigène-anticorps formé.
- 50 12. Test immunologique selon la revendication 11 dans lequel l'échantillon de test comprend du sang humain ou une fraction de celui-ci.
- 55

Revendications pour l'Etat contractant suivant : ES

1. Méthode de préparation d'un polynucléotide sous une forme sensiblement isolée qui comprend une séquence nucléotidique d'au moins 15 nucléotides d'un isolat HCV J-1, ledit isolat HCV J-1 ayant une homologie de séquence nucléotidique d'au moins 90 % avec la séquence J-1 de l'une quelconque des figures 7 à 10 ou 13 à 18, ladite séquence nucléotidique d'au moins 15 nucléotides étant différente de la séquence nucléotidique de l'isolat HCV-1 de HCV, ladite méthode comprenant :
  - a) la synthèse chimique par des méthodes connues *per se* ; ou
  - b) la réplication d'ADN par des méthodes connues *per se* ; ou
  - c) la transcription ou rétrotranscription par des méthodes connues *per se* ; ou
  - d) la digestion par des enzymes de restriction et la ligature dans un vecteur par des techniques de l'ADN recombinant connues *per se*.
2. Méthode de préparation d'un polypeptide comprenant une séquence d'acides aminés qui est codée par une séquence nucléotidique telle que définie à la revendication 1, ladite méthode comprenant :
  - (i) la synthèse chimique par des méthodes connues *per se* ; ou
  - (ii) la traduction à partir de la séquence d'acide nucléique correspondante par des méthodes connues *per se* ; ou
  - (iii) l'expression d'un système d'expression recombinant comprenant la séquence nucléotidique codant pour le polypeptide par des méthodes connues *per se* ; ou
  - (iv) l'isolement à partir de virus par des méthodes connues *per se*.
3. Méthode de détection de polynucléotides de HCV dans un échantillon de test comprenant :
  - (a) la fourniture d'un polynucléotide tel que défini à la revendications 1 comme sonde ;
  - (b) la mise en contact de l'échantillon de test et de la sonde dans des conditions qui permettent la formation d'un duplex polynucléotidique entre la sonde et son complément en l'absence de formation appréciable de duplex polynucléotidique entre la sonde et des séquences polynucléotidiques autres que de HCV présentes dans l'échantillon de test ; et
  - (c) la détection de tous duplex polynucléotidiques comprenant la sonde.
4. Test immunologique destiné à détecter la présence d'anticorps anti-HCV dans un échantillon de test qui comprend :
  - (a) l'incubation de l'échantillon de test dans des conditions qui permettent la formation d'un complexe antigène-anticorps à former avec, un polypeptide qui est codé par une séquence nucléotidique telle que définie à la revendication 1 ou une séquence d'au moins 15 nucléotides d'un isolat HCV J-1 présent dans l'un quelconque des plasmides pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW 770-BP-N ou J1 5-1-1 déposés sous les numéros d'accès BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 et 40884 respectivement, ladite séquence nucléotidique d'au moins 15 nucléotides étant différente de la séquence nucléotidique de l'isolat HCV-1 de HCV, ladite séquence codante étant en phase avec les séquences d'acides aminés correspondantes représentées dans les figures 7 à 10 ou 13 à 18, lequel polypeptide comprend un déterminant antigénique et est distinct de la séquence du polypeptide codé par l'isolat HCV-1 de HCV et ne produit pas de réaction immunologique croisée avec HCV-1 ; et
  - (b) la détection de tous complexes antigène-anticorps formés.
5. Test immunologique selon la revendication 6 dans lequel l'échantillon de test comprend du sang humain ou une fraction de celui-ci.
6. Polynucléotide sous une forme sensiblement isolée comprenant une séquence nucléotidique d'au moins 15 nucléotides d'un isolat HCV J-1, ledit isolat HCV J-1 ayant une homologie de séquence nucléotidique d'au moins 90 % avec la séquence J-1 de l'une quelconque des figures 7 à 10 ou 13 à 18, dans lequel ladite séquence nucléotidique d'au moins 15 nucléotides est différente de la séquence nucléotidique de l'isolat HCV-1 de HCV.
7. Polynucléotide selon la revendication 6 dans lequel l'isolat HCV J-1 a une homologie d'au moins 95 % avec la séquence de HCV J-1 de l'une quelconque des figures 7 à 10 ou 13 à 18.

**EP 0 419 182 B1**

8. Polynucléotide selon la revendication 6 dans lequel l'isolat HCV J-1 a une homologie de 100 % avec la séquence de HCV J-1 de l'une quelconque des figures 7 à 10 ou 13 à 18.
9. Polynucléotide selon l'une quelconque des revendications 6 à 8 qui comprend au moins 20 nucléotides.
10. Polynucléotide comprenant une séquence d'au moins 15 nucléotides d'un isolat HCV J-1 présent dans l'un quelconque des plasmides pS1-8791a, bU1-1216c, bU1-4652d, pS1-713c, pS1-1519, TC-600BP, JH-400BP, AW-300BP, AW 770-BP-N ou J1 5-1-1 déposés sous les numéros d'accès BP-2593, BP-2594, BP-2595, BP-2637, BP-3081, 68393, 68394, 68392, 68395 et 40884 respectivement, ladite séquence nucléotidique d'au moins 15 nucléotides étant différente de la séquence nucléotidique de l'isolat HCV-1 de HCV.
11. Polypeptide purifié comprenant une séquence d'acides aminés qui :
- (a) est codée par une séquence nucléotidique telle que définie dans l'une quelconque des revendications 6 à 9 ou dans les séquences de HCV déposées et définies à la revendication 10, ladite séquence codante étant en phase avec les séquences d'acides aminés correspondantes présentées dans les figures 7 à 10 ou 13 à 18 ;
  - (b) comprend un déterminant antigénique ; et
  - (c) est distincte de la séquence du polypeptide codé par l'isolat HCV-1 de HCV.
12. Polypeptide selon la revendication 11 qui comprend au moins 10 acides aminés.
13. Polypeptide selon la revendication 11 qui comprend au moins 15 acides aminés.
14. Polypeptide selon l'une quelconque des revendications 11 à 13 immobilisé sur un support solide.



J7 1 AGCCGAGTAGTGTGGGTGCGGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTTGGCAG 60  
discrepancy  
clone  
altered aa

J7 61 TCCCCGGGAGGTCTCGTAGACCGTGCATCATGAGCACAAATCCTAAACCCCAAGAAAA 120  
MetSerThrAsnProLysProGlnArgLys  
TG  
b1  
---Arg

J7 121 ACCAAACGTAAACACCAACCGTCGCCCACAGGACGTTAAGTTCCTGGGCGGTGGTCAGATC 180  
ThrLysArgAsnThrAsnArgArgProGlnAspValLysPheProGlyGlyGlyGlnIle  
C T  
b 1  
--- Leu

J7 181 GTCGGTGGAGTTTACTTGTTCGCGCGCAGGGGCCCCAGGTTGGGTGTGCGTGGCACTAGG 240  
ValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeuGlyValArgAlaThrArg  
T A  
b b  
--- ---

J7 241 AAGACTTCCGAGCGGTGCGAACCTCGTGGAAGGCGACAACCTATCCCCAAGGTCGCGCG 300  
LysThrSerGluArgSerGlnProArgGlyArgArgGlnProIleProLysAlaArgArg  
A  
b  
---

J7 301 CCCGAGGGCAGGACCTGGGCTCAGCCTGGGTATCCTTGGCCCCCTCTATGGCAATGAGGGC 360  
ProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpProLeuTyrGlyAsnGluGly

J7 361 TTGGGGTGGCAGGATGGCTCCTGTACCCCCGGGCTCTCGGCCTAGTTGGGGCCCCAAT 420  
LeuGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArgProSerTrpGlyProAsn  
A T C  
b c b  
END ---Thr

J7 421 GACCCCGGCGTAGGTCGCGTAATTTGGGTAAGGTCATCGATACCTTACATGCGGCTTC 480  
AspProArgArgArgSerArgAsnLeuGlyLysValIleAspThrLeuThrCysGlyPhe  
C  
1  
Leu

FIGURE 1-1

EP 0 419 182 B1

		AlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAla	
J7	481	GCCGACCTCATGGGGTACATTCCGCTTGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCC	540
		<div style="display: flex; justify-content: space-around; width: 100%;"> <span>C</span> <span>C</span> </div> <div style="display: flex; justify-content: space-around; width: 100%;"> <span>c</span> <span>b</span> </div> <div style="display: flex; justify-content: space-around; width: 100%;"> <span>---</span> <span>---</span> </div>	
		LeuAlaHisGly	
J7	541	CTGGCACATGGT	552

FIGURE 1-2

J1 1 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 61  
 discrepancy TCCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCACATGGTGTCCGGGTT  
 clone C  
 altered aa d Ser

62 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 121  
 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTCTATCTTC

122 LeuLeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValArgAsnVal 181  
 CTCTGGCTCTGCTGTCCTGTTTGACCATCCCAGCTTCCGCTTATGAAGTGCGCAACGTG  
 A T  
 g d  
 --- ---

182 SerGlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAla 241  
 TCCGGGATATACCATGTCACAAACGACTGCTCCAACTCAAGCATTGTGTATGAGGCGGCG  
 T  
 d  
 ---

242 AspValIleMetHisAlaProGlyCysValProCysValArgGluAsnAsnSerSerArg 301  
 GACGTGATCATGCATGCCCCGGGTGCGTGCCCTGCGTTCGGGAGAACAAATTCCTCCCGT  
 C  
 d  
 ---

302 CysTrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrThr 361  
 TGCTGGGTAGCGCTCACTCCCACGCTCGCGGCCAGGAATGCCAGCGTCCCCACTACGACA

362 LeuArgArgHisValAspLeuLeuValGlyThrAlaAlaPheCysSerAlaMetTyrVal 421  
 TTACGACGCCACGTCGACTTGCTCGTTGGGACGGCTGCTTTCTGCTCCGCTATGTACGTG  
 G  
 d  
 ---

FIGURE 2-1

422	<p>GlyAspLeuCysGlySerValPheLeuIleSerGlnLeuPheThrPheSerProArgArg</p> <p>GGGGATCTCTGCGGATCTGTTTTCTCATCTCCAGCTGTTACCTTCTCGCCTCGCCGG</p>	481
	<p>T</p> <p>d</p> <p>---</p>	
482	<p>HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArg</p> <p>CATGAGACAGTACAGGACTGCAACTGCTCAATCTATCCCGGCCACGTATCAGGCCATCGC</p>	541
	<p>T</p> <p>C</p> <p>---</p>	
542	<p>MetAlaTrpAspMetMetMetAsnTrpSerProThrAla</p> <p>ATGGCTTGGGATATGATGATGAACTGGTCGCCCACGGCA</p>	580

FIGURE 2-2

			AsnTrpSerProThrAla	
J1	1	AACTGGTCGCCCACGGCA	18	
		discrepancy	!!!	
		clone		
		altered aa		
J1	19	AlaLeuValValSerGlnLeuLeuArgIleProGlnAlaValMetAspMetValAlaGly GCCTTAGTGGTGTCCGAGTTACTCCGGATCCACAAGCTGTCATGGACATGGTGGCGGGG	78	
J1	79	AlaHisTrpGlyValLeuAlaGlyLeuAlaTyrTyrSerMetValGlyAsnTrpAlaLys GCCCCACTGGGGAGTCCTAGCGGGCCTTGCTACTATTCCATGGTGGGGAAGTGGGCTAAG	138	
		G A		
		i i		
		---	---	
J1	139	ValLeuIleValMetLeuLeuPheAlaGlyValAspGlyHisThrArgValThrGlyGly GTTTTGATTGTGATGCTACTCTTTGCCGGCGTTGACGGGCATACCCGCGTGACGGGGGGG	198	
		AG A		
		GG i		
		Ser ---		
J1	199	ValGlnGlyHisValThrSerThrLeuThrSerLeuPheArgProGlyAlaSerGlnLys GTGCAAGGCCACGTCACCTCTACACTCAGTCCCTCTTTAGACCTGGGGCGTCCCAGAAA	258	
		T G		
		c i		
		---	Ala	
J1	259	IleGlnLeuValAsnThrAsnGlySerTrpHisIleAsnArgThrAlaLeuAsnCysAsn ATTCAGCTTGTAACACCAATGGCAGTTGGCATATCAACAGGACTGCCCTGAACTGCAAT	318	
		TC T		
		ii i		
		SerLeu		
			---	
J1	319	AspSerLeuGlnThrGlyPheLeuAlaAlaLeu GACTCCCTCCAAACTGGGTTCTTGCCGCGCTG	351	

FIGURE 3

				J1	1	Ser CTCA	4
						discrepancy clone altered aa	
J1	5	ValIleAspCysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPhe GTGATCGACTGTAACACATGTGTCACTCAGACGGTCGATTTCAGCTTGGATCCCACCTTC	64			G C Ala	
J1	65	ThrIleGluThrThrThrValProGlnAspAlaValSerArgThrGlnArgArgGlyArg ACCATCGAGACGACGACCGTGCCCAAGATGCGGTTTCGCGCACGCAGCGCGGAGGTAGG	124				
J1	125	ThrGlyArgGlyArgArgGlyIleTyrArgPheValThrProGlyGluArgProSerAla ACTGGCAGGGGCAGGAGAGGCATCTATAGGTTTGTGACTCCAGGAGAACGGCCCTCGGCG	184				
J1	185	MetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeu ATGTTTCGATTCTTCGGTCCTATGTGAGTGTTATGACCGGGCTGTGCTTGGTATGAGCTC	244			A Gly(-)	
J1	245	ThrProAlaGluThrSerValArgLeuArgAlaTyrLeuAsnThrProGlyLeuProVal ACGCCGCTGAGACCTCGGTTAGGTTGCGGGCTTACCTAAATACACCAGGGTTGCCCGTC	304				
J1	305	CysGlnAspHisLeuGluPheTrpGluSerValPheThrGlyLeuThrHisIleAspAla TGCCAGGACCATCTGGAGTTCCTGGGAGAGCGTCTTCACAGGCCTCACCCACATAGACGCC	364				
J1	365	HisPheLeuSerGlnThrLysGlnAlaGlyAspAsnPheProTyrLeuValAlaTyrGln CACTTCTTGTCAGACTAAGCAGGCAGGAGACAACCTCCCTACCTGGTAGCATACCA	424				

FIGURE 4-1

;

				LeuThr	
	J1		1	CCTCACC	7
		discrepancy			
		clone			
		altered aa			
J1	8	ArgAspProThrValProLeuAlaArgAlaAlaTrpGluThrAlaArgHisThrProVal			
		CGTGACCCACCGTCCCCCTTGCGGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTC	67		
				C	
				G	
				Thr(=)	
J1	68	AsnSerTrpLeuGlyAsnIleIleMetTyrAlaProThrLeuTrpAlaArgMetIleLeu			
		AACCTCTGGCTAGGCAACATCATCATGTATGCGCCCACTTTGTGGGCAAGGATGATTCTG	127		
				T	
				G	
				Ile(=)	
J1	128	MetThrHisPhePheSerIleLeuLeuAlaGlnGluGlnLeuGluLysAlaLeuAspCys			
		ATGACTCACTTCTTCTCCATCCTTCTAGCCAGGAGCAACTTGAAAAAGCCCTGGATTGT	187		
J1	188	GlnIleTyrGlyAlaCysTyrSerIleGluProLeuAspLeuProGlnIleIleGluArg			
		CAAATCTACGGGGCCTGTTACTCCATTGAGCCACTTGACCTACCTCAGATCATTGAACGA	247		
J1	248	LeuHisGlyLeuSerAlaPheSerLeuHisSerTyrSerProGlyGluIleAsnArgVal			
		CTCCATGGTCTTAGCGCATTTCCTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTG	307		
J1	308	AlaSerCysLeuArgLysLeuGlyValProProLeuArgValTrpArgHisArgAlaArg			
		GCTTCATGCCTCAGGAAGCTTGGGTACCACCCTTGCGAGTCTGGAGACATCGGGCCAGA	367		
J1	368	SerValArgAlaLysLeuLeuSerGlnGlyGlyArgAlaAlaThrLysGlyLysTyrLeu			
		AGTGTCCGCGCTAAGCTACTGTCCCAAGGGGGAGGGCCGCACTTGTGGCAAGTACCTC	427		
				G	
				G	
				Gln(=)	

FIGURE 5



J7 1 AGCCGAGTAGTGTGGGTGCGAAAGGCCCTTGTGGTACTGCCTGATAGGGTGCTTGCGAG 60  
HCV1

J7 61 TCCCCCGGGAGGTCTCGTAGACCGTGCATCATGAGCACAATCCTAAACCCCAAAGAAAA 120  
HCV1 MetSerThrAsnProLysProGlnArgLys  
C G T A  
Lys

J7 121 ThrLysArgAsnThrAsnArgArgProGlnAspValLysPheProGlyGlyGlyGlnIle 180  
HCV1 ACCAAACGTAAACACCAACCGTCGCCACAGGACGTTAAGTCCCGGGCGGTGCTCAGATC  
A C T C  
Asn  
...

J7 181 ValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeuGlyValArgAlaThrArg 240  
HCV1 GTCGGTGGAGTTTACTTGTGCCGCGCAGGGCCCCAGGTTGGGTGTGCGTGCGACTAGG  
T T A C G A

J7 241 LysThrSerGluArgSerGlnProArgGlyArgArgGlnProIleProLysAlaArgArg 300  
HCV1 AAGACTTCCGAGCGGTGCGAACCTCGTGAAGGCGACAACCTATCCCCAAGGCTCGCCGG  
A T A T G T

J7 301 ProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpProLeuTyrGlyAsnGluGly 360  
HCV1 CCCGAGGGCAGGACCTGGGCTCAGCCTGGGTATCCTTGGCCCCCTCTATGGCAATGAGGGC  
C C

J7 361 LeuGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArgProSerTrpGlyProAsn 420  
HCV1 TTGGGGTGGGCAGGATGGCTCCTGTACCCCGCGGCTCTCGGCCTAGTTGGGGCCCCAAT  
GC G T T C CA  
Cys Thr  
...

J7 421 AspProArgArgArgSerArgAsnLeuGlyLysValIleAspThrLeuThrCysGlyPhe 480  
HCV1 GACCCCCGCGTAGGTGCGTAATTTGGGTAAGGTCATCGATACCCTTACATGCGGCTTC  
C G

FIGURE 6-1

EP 0 419 182 B1

		AlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAla	
J7	481	GCCGACCTCATGGGTACATTCCGCTTGTGGCGCCCCCTTAGGGGGCGCTGCCAGGGCC	540
HCV1		A C TC T A	
		LeuAlaHisGly	
J7	541	CTGGCACATGGT	552
HCV1		G C	781

FIGURE 6-2

J1 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 61  
 HCV1 1 TCCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCACATGGTGTCCGGTT  
 A TC T A G C

J1 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 121  
 HCV1 62 CTGGAGGACGGCGTGAACATATGCAACAGGGAATTGCCCCGGTTGCTCTTCTCTATCTTC  
 A CC T T

J1 LeuLeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValArgAsnVal 181  
 HCV1 122 CTCTTGGCTCTGCTGTCTGTTTGACCATCCAGCTTCCGTTATGAAGTGCAGCAACGTG  
 TC C C T C TG G C G C CC TCC  
 Val Gln Ser  
 \*\*\* \*\*\*

J1 SerGlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAla 241  
 HCV1 182 TCCGGGATATACCATGTACAAACGACTGCTCCAACCAAGCATTGTGTATGAGGCGCG  
 A G C T C C T T C T G T C C  
 Thr Leu Pro

J1 AspValIleMetHisAlaProGlyCysValProCysValArgGluAsnAsnSerSerArg 301  
 HCV1 242 GACGTGATCATGCATGCCCGGGGTGCGTCCCTGCGTTCCGGAGAACAAATTCCTCCCGT  
 T CC C CA T G C T T GG CG GA G  
 Ala Leu Thr Gly Ala  
 \*\*\*

J1 CysTrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrThr 361  
 HCV1 302 TGCTGGGTAGCGCTCACTCCACGCTCGCGGCCAGGAATGCCAGCGTCCCACTACGACA  
 T G A G C T G G CA G G AAC G G CAG  
 Met Val Thr AspGlyLysLeu Ala Gln  
 \*\*\* \*\*\*

J1 LeuArgArgHisValAspLeuLeuValGlyThrAlaAlaPheCysSerAlaMetTyrVal 421  
 HCV1 362 TTACGACGCCACGTCGACTTGCTCGTTGGGACGGCTGCTTTCTGCTCCGCTATGTACGTG  
 C T T A TC T C GC CA CC T G CC C  
 Ile Ser ThrLeu Leu

J1 GlyAspLeuCysGlySerValPheLeuIleSerGlnLeuPheThrPheSerProArgArg 481  
 HCV1 422 GGGGATCTCTGCGGATCTGTTTTCTCATCTCCAGCTGTTACCTTCTCGCCTCGCCGG  
 C A G C T TG GG A T CA G C  
 ValGly

FIGURE 7-1

J1	482	HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArg	
HCV1		CATGAGACAGTACAGGACTGCAACTGCTCAATCTATCCCGGCCACGTATCAGGCCATCGC	341
		CTG GACG A GT T T TA A G T C	
		Trp Thr Gly IleThr	
		... ...	
J1	542	MetAlaTrpAspMetMetMetAsnTrpSerProThrAla	
HCV1		ATGGCTTGGGATATGATGATGAACTGGTCGCCCACGGCA	580
		A C T A G	
		Thr	

FIGURE 7-2

					AsnTrpSerProThrAla	
J1				1	AAGTGGTCGCCCCACGGCA	18
HCV1					C T A G	
					Thr	
J1	19	AlaLeuValValSerGlnLeuLeuArgIleProGlnAlaValMetAspMetValAlaGly				78
HCV1		GCCTTAGTGGTGTCGCAGTTACTCCGGATCCCAAGCTGTCATGGACATGGTGGCGGGG				
		G G A A G T C G		CA T	A C T T	
		MetAla		IleLeu	Ile	
J1	79	AlaHisTrpGlyValLeuAlaGlyLeuAlaTyrTyrSerMetValGlyAsnTrpAlaLys				138
HCV1		GCCCACTGGGGAGTCTAGCGGGCCTTGCCCTACTATTCCATGGTGGGAACTGGGCTAAG				
		T G A A G T TC			G	
		Ile Phe				
J1	139	ValLeuIleValMetLeuLeuPheAlaGlyValAspGlyHisThrArgValThrGlyGly				198
HCV1		GTTTTGATTGTGATGCTACTCTTTGCCGGCGTTGACGGGCATACCCGCGTGACGGGGGGG				
		CC G A C G A C C G A A C C A				
		Val Leu		AlaGlu His		
				...		
J1	199	ValGlnGlyHisValThrSerThrLeuThrSerLeuPheArgProGlyAlaSerGlnLys				258
HCV1		GTGCAAGGCCACGTCACCTCTACACTCAAGTCCCTCTTTAGACCTGGGGCGTCCCAGAAA				
		AGTGCC ACTGTG GG T TGTTAG C CGC A C CAAG C				
		SerAla ThrVal GlyPheVal LeuAla Lys Asn				
		.....		...	...	
J1	259	IleGlnLeuValAsnThrAsnGlySerTrpHisIleAsnArgThrAlaLeuAsnCysAsn				318
HCV1		ATTGAGCTTGTAACACCAATGGCAGTTGGCATATCAACAGGACTGCCCTGAACTGCAAT				
		G C GA C C CC T C G				
		Val Ile		Leu Ser		
				...		
J1	319	AspSerLeuGlnThrGlyPheLeuAlaAlaLeu				351
HCV1		GACTCCCTCCAACTGGGTTCTTGCCGCGCTG				
		TAG A C C C GGT G A G T				
		Asn Trp Gly				

FIGURE 8

J1 1 SerValIleAspCysAsnThrCysValThrGlnThrVal  
 HCV1 1 C T C A G T G C A  
 55ctataccggcgacttcga G A C T G C A  
  
 J1 41 AspPheSerLeuAspProThrPheThrIleGluThrThrThrValProGlnAspAlaVal  
 HCV1 41 GATTTCAGCTTGGATCCACCTTCACCATCGAGACGACGACCGTGCCCAAGATGCGGT  
 C T C T T A T C G C C G T C  
 Ile  
  
 J1 101 SArArgThrGlnArgArgGlyArgGlyArgGlyIleTyrArgPheVal  
 HCV1 101 TCGCGCACGCAGCGCGGAGGTAGGACTGGCAGGGGCAGGAGGCATCTATAGGTTTGT  
 C T A T G C G A C C C A  
 LysPro  
  
 J1 181 ThrProGlyGluArgProSerAlaMetPheAspSerSerValLeuCysGluCysTyrAsp  
 HCV1 181 ACTCCAGGAGAACGGCCCTCGGCCATGTTTCGATTCTTCGGTCTATGTGAGTGTATGAC  
 G A G G G C C G C C G C C C C  
 Ala Gly  
  
 J1 221 AlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThrSerValArgLeuArgAlaTyr  
 HCV1 221 GCGGGCTGTGCTTGGTATGAGCTCACGCCGCTGAGACCTCGGTTAGGTTGCGGGCTTAC  
 A C T A A C A A G  
 Thr  
  
 J1 281 LeuAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluSerValPhe  
 HCV1 281 C T A A T A C A C C A G G G T T G C C C G T C T G C C A G G A C C A T C T G G A G T T C T G G A G A G C G T C T T C  
 A G C C G C T G T A T G G Lys  
 Met  
  
 J1 341 ThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnAlaGlyAspAsn  
 HCV1 341 A C A G G C C T C A C C C A C A T A G A C G C C C A C T T C T G T C C C A G A C T A A G C A G G C A G G A G A C A A C  
 T T T T C A A A G T G G  
 Ser Glu  
  
 J1 401 PheProTyrLeuValAlaTyrGlnAlaThrValCysAlaArgAlaLysAlaProProPro  
 HCV1 401 T T C C C C T A C C T G G T A G C A T A C C A A G C C A C A G T G T G C G C C A G G G C T A A G G C T C C A C C T C C A  
 C T T G C T C A C T C  
 Leu GLE  
  
 J1 481 SerTrpAspGlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThr  
 HCV1 481 T C G T G G G A T C A A A T G T G G A A G T G T C T C A T A C G G C T A A A G C C T A C G C T G C A C G G G G C C A A C C  
 G G T G T C C C C C C C T  
  
 J1 521 ProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProIleThr  
 HCV1 521 C C C C T G C T G T A T A G G C T A G G A G C C G T C C A G A A T G A G G T C A C C C T C A C A C C C T A T A A C C  
 A C A G C T T A A G G C  
 Ile Val  
  
 J1 581 Lys  
 HCV1 581 AAA  
 tacatcatgacatgcatgtc

FIGURE 9

LeuThr  
J1 1 CCTCACC 7  
HCV1

ArgAspProThrValProLeuAlaArgAlaAlaTrpGluThrAlaArgHisThrProVal  
J1 8 CGTGACCCACCGTCCCCCTTGC GCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTC 67  
HCV1 T AAC C A A A  
Thr  
...

AsnSerTrpLeuGlyAsnIleIleMetTyrAlaProThrLeuTrpAlaArgMetIleLeu  
J1 68 AACTCCTGGCTAGGCAACATCATCATGTATGCGCCCACTTTGTGGGCAAGGATGATTCTG 127  
HCV1 T A T C AC G A  
Phe

MetThrHisPhePheSerIleLeuLeuAlaGlnGluGlnLeuGluLysAlaLeuAspCys  
J1 128 ATGACTCACTTCTTCTCCATCCTTCTAGCCCAGGAGCAACTTGAAAAAGCCCTGGATTGT 187  
HCV1 C T TAG G A AG C G C G C C  
Val Ile ArgAsp Gln  
...

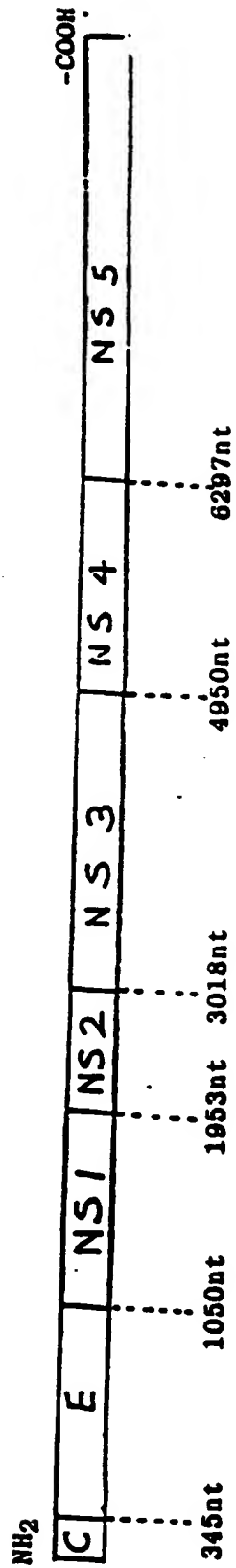
GlnIleTyrGlyAlaCysTyrSerIleGluProLeuAspLeuProGlnIleIleGluArg  
J1 188 CAAATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTCAGATCATTGAACGA 247  
HCV1 G G C A A T CA C A  
Glu Pro Gln  
...

LeuHisGlyLeuSerAlaPheSerLeuHisSerTyrSerProGlyGluIleAsnArgVal  
J1 248 CTCCATGGTCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTG 307  
HCV1 C C C A T

AlaSerCysLeuArgLysLeuGlyValProProLeuArgValTrpArgHisArgAlaArg  
J1 308 GCTTCATGCCTCAGGAAGCTTGGGGTACCACCCTTGGGAGTCTGGAGACATCGGGCCAGA 367  
HCV1 CG A A G CT C C G  
Ala Ala  
...

SerValArgAlaLysLeuLeuSerGlnGlyGlyArgAlaAlaThrLysGlyLysTyrLeu  
J1 368 AGTGTCCGCGCTAAGCTACTGTCCCAAGGGGGAGGGCCGCACTTGTGGCAAGTACCTC 427  
HCV1 C G T G AG A C T TA  
Arg AlaArg Ile  
...

FIGURE 10



The nucleotide numbers are approximate.

FIGURE 11



## HCV-1

-367 GCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTCAGCCTCCAGGACCCCCCTCCGGGA  
CGCAGATCGGTACCGCAATCATACTCACAGCACGTCTGGAGGTCTGGGGGGGAGGGCCCT

-207 GAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGACCGGGTCTCT  
CTCGGTATCACCAGACGCCTTGGCCACTCATGTGGCCTTAACGGTCTGTGCGCCAGGA

-147 TTCTTGGATCAACCCGCTCAATGCCTGGAGATTGGGCGTGCCCCGCAAGACTGCTAGC  
AAGAACCCTAGTTGGGCGAGTTACGGACCTCTAAACCCGCACGGGGCGTTCTGACGATCG

-87 CGAGTAGTGTGGGTCTGCGAAAGGCCTTGTGGTACTGCCTGATAGGGTGTCTTGGCAGTGC  
GCTCATCACAACCCAGCGCTTTCGGAACACCATGACGGACTATCCACGAACGCTCAGC

-27 CCCGGGAGGTCTCGTAGACCGTGCACC -1  
GGGCCCTCCAGAGCATCTGGCACGTGG

MetSerThrAsnProLysProGlnLysLysAsnLysArgAsnThrAsnArgArgProGln  
1 ATGAGCACGAATCCTAAACCTCAAAAAAAAAAAACAAACGTAACACCAACCGTCGCCACAG  
TACTCGTGCTTAGGATTGGAGTTTTTTTTTTGTTCATTGTGGTTGGCAGCGGGTGTC

AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArg  
61 GAGCTCAAGTTCCCGGGTGGCGGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT  
CTGCAGTTCAAGGGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGGTCC

GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly  
121 GGCCCTAGATTGGGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTTCGCAACCTCGAGGT  
CCGGGATCTAACCACACGCGCGTCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCA

ArgArgGlnProIleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGly  
181 AGACGTCAGCCTATCCCCAAGGCTCTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGG  
TCTGCAGTCGGATAGGGGTTCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCTGGGGCC

TyrProTrpProLeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerPro  
241 TACCCTTGGCCCCCTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCTGTCTCCC  
ATGGGAACCGGGGAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGG

ArgGlySerArgProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGly  
301 CGTGGCTCTCGGCCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTTCGCGCAATTTGGGT  
GCACCGAGAGCCGGATCGACCCGGGGTGTCTGGGGGCCGATCCAGCGCGTTAAACCCA

LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuVal  
361 AAGGTCATCGATACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTC  
TTCCAGTAGCTATGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCAGCAG

GlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp  
421 GGCGCCCCCTCTGGAGGGCGTCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGAC  
CCGCGGGGAGAACCTCCGCGACGGTCCCGGGACCGGTACCGCAGGCCCAAGACCTTCTG

GlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla  
481 GGCGTGAACCTATGCAACAGGAACCTTCTGGTTGCTCTTCTCTATCTTCTTCTGGCC  
CCGCACCTTGATACGTTGTCCCTTGAAGGACCAACGAGAAAGAGATAGAAGGAAGACCGG

FIGURE 12-1

541 LeuLeuSerCysLeuThrValProAlaSerAlaTyrGlnValArgAsnSerThrGlyLeu  
 CTGCTCTCTTGCTTGACTGTGCCCGCTTCGGCCTACCAAGTGCAGCACTCCACGGGGCTT  
 GACGAGAGAACGAACTGACACGGGGCGAAGCCGGATGGTTACGCGTTGAGGTGCCCGAA  
  
 601 TyrHisValThrAsnAspCysProAsnSerSerIleValTyrGluAlaAlaAspAlaIle  
 TACCACGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGGCGGCCGATGCCATC  
 ATGGTGCAGTGGTTACTAACGGGATTGAGCTCATAACACATGCTCCGCCGCTACGGTAG  
  
 661 LeuHisThrProGlyCysValProCysValArgGluGlyAsnAlaSerArgCysTrpVal  
 CTGCACACTCCGGGGTGCCTCCCTTTCGTTTCGTGAGGGCAACGCCTCGAGGTGTGGGTG  
 GACGTGTGAGGCCCCACGCAGGGAACGCAAGCACTCCCGTTGCGGAGCTCCACAACCCAC  
  
 721 AlaMetThrProThrValAlaThrArgAspGlyLysLeuProAlaThrGlnLeuArgArg  
 GCGATGACCCCTACGGTGGCCACCAGGGATGGCAACTCCCCGCGACGCAGCTTCGACGT  
 CGCTACTGGGGATGCCACCGTGGTCCCTACCGTTTGAGGGGCGCTGCGTCGAAGCTGCA  
  
 781 HisIleAspLeuLeuValGlySerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu  
 CACATCGATCTGCTTGTTCGGGAGCGCCACCCTCTGTTCCGCCCTCTACGTGGGGGACCTA  
 GTGTAGCTAGACGAACAGCCCTCGCGGTGGGAGACAAGCCGGGAGATGCACCCCTGGAT  
  
 841 CysGlySerValPheLeuValGlyGlnLeuPheThrPheSerProArgArgHisTrpThr  
 TCGGGTCTGTCTTTCTTGTTCGCCCACTGTTACCTTCTCTCCAGGCGCCACTGGACG  
 ACGCCAGACAGAAAGAACAGCCGGTTGACAAGTGAAGAGAGGGTCCGCGGTGACCTGC  
  
 901 ThrGlnGlyCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArgMetAlaTrp  
 ACGCAAGGTGCAATTGCTCTATCTATCCCGGCCATATAACGGGTCACCGCATGGCATGG  
 TCGGTTCCAACGTTAACGAGATAGATAGGGCCGGTATATTGCCAGTGGCGTACCGTACC  
  
 961 AspMetMetMetAsnTrpSerProThrThrAlaLeuValMetAlaGlnLeuLeuArgIle  
 GATATGATGATGAAGTGGTCCCTACGACGGCGTTGGTAATGGCTCAGCTGCTCCGGATC  
 CTATACTACTACTTGACCAGGGGATGCTGCCGCAACCATTACCGAGTCGACGAGGCCTAG  
  
 1021 ProGlnAlaIleLeuAspMetIleAlaGlyAlaHisTrpGlyValLeuAlaGlyIleAla  
 CCACAAGCCATCTTGGACATGATCGCTCGTGCTCACTGGGGAGTCTGGCGGGCATAGCG  
 GGTGTTCCGGTAGAACCTGTACTAGCGACCAGAGTGACCCCTCAGGACCGCCCGTATCGC  
  
 1081 TyrPheSerMetValGlyAsnTrpAlaLysValLeuValValLeuLeuPheAlaGly  
 TATTTCTCCATGGTGGGGAAGTGGGCGAAGGTCTCGGTAGTGCTGCTGCTATTTGCCGGC  
 ATAAAGAGGTACCAACCCCTTGACCCGCTTCAGGACCATCACGACGACGATAAACGGCCG  
  
 1141 ValAspAlaGluThrHisValThrGlyGlySerAlaGlyHisThrValSerGlyPheVal  
 GTCGACGCGGAAACCCACGTACCGGGGGAAGTGCCGGCCACACTGTGTCTGGATTGTGT  
 CAGCTGCGCCTTTGGGTGCAGTGGCCCCCTTCACGGCCGGTGTGACACAGACCTAAACAA  
  
 1201 SerLeuLeuAlaProGlyAlaLysGlnAsnValGlnLeuIleAsnThrAsnGlySerTrp  
 AGCCTCCTCGCACCAGGCGCAAGCAGAAGCTCCAGCTGATCAACACCAACGGCAGTTGG  
 TCGGAGGAGCGTGGTCCGCGGTTCTGCTTGCAGGTGCGACTAGTTGTGGTTGCCGTCAACC  
  
 1261 HisLeuAsnSerThrAlaLeuAsnCysAsnAspSerLeuAsnThrGlyTrpLeuAlaGly  
 CACCTCAATAGCACGGCCCTGAAGTGAATGATAGCCTCAACACCGGCTGGTTGGCAGGG  
 GTGGAGTTATCGTGCCGGGACTTGACGTTACTATCGGAGTTGTGGCCGACCAACCGTCCC  
  
 1321 LeuPheTyrHisHisLysPheAsnSerSerGlyCysProGluArgLeuAlaSerCysArg  
 CTTTTCTATCACCACAAGTTCAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGA  
 GAAAAGATAGTGGTGTCAAGTTGAGAAGTCCGACAGGACTCTCCGATCGGTCCAGCGCT

FIGURE 12-2

1381 ProLeuThrAspPheAspGlnGlyTrpGlyProIleSerTyrAlaAsnGlySerGlyPro  
 CCCCTTACCGATTTTGACCAGGGCTGGGGCCCTATCAGTTAIGCCAACGGAAGCGGCCCC  
 GGGGAATGGCTAAAACTGGTCCCACCCCGGGATAGTCAATACGTTGCCTTCGCCGGGG  
 1441 AspGlnArgProTyrCysTrpHisTyrProProLysProCysGlyIleValProAlaLys  
 GACCAGCGCCCTACTGCTGGCACTACCCCCAAAACCTTGGGATTTGTGCCCCGGAAG  
 CTGGTCGCGGGGATGACGACCGTGATGGGGGTTTTGGAACGCCATAACACGGGCGCTTC  
 1501 SerValCysGlyProValTyrCysPheThrProSerProValValValGlyThrThrAsp  
 AGTGTGTGTGGTCCGGTATATTGCTTCACTCCAGCCCCGTGGTGGTGGGAACGACCGAC  
 TCACACACACCAGGCCATATAACGAAGTGAGGGTCGGGGCACCAACACCTTGCTGGCTG  
 1561 ArgSerGlyAlaProThrTyrSerTrpGlyGluAsnAspThrAspValPheValLeuAsn  
 AGGTCGGGCGCGCCACCTACAGCTGGGGTGAAAATGATACGGACGTCTTCGTCTTAAC  
 TCCAGCCCGCGCGGTGGATGTCGACCCCACTTTTACTATGCCTGCAGAAGCAGGAATTG  
 1621 AsnThrArgProProLeuGlyAsnTrpPheGlyCysThrTrpMetAsnSerThrGlyPhe  
 AATACCAGGCCACCGCTGGGCAATTGGTTCGGTTGTACCTGGATGAACCTCAACTGGATTTC  
 TTATGGTCCGGTGGCGACCCGTTAACCAAGCCAACATGGACCTACTTGAGTTGACCTAAG  
 1681 ThrLysValCysGlyAlaProProCysValIleGlyGlyAlaGlyAsnAsnThrLeuHis  
 ACCAAAGTGTGCGGAGCGCTCCTTGTGTTCATCGGAGGGGCGGGCAACAACACCTGCAC  
 TGGTTTTACACGCCTCGCGGAGGAACACAGTAGCCTCCCCCGCGTTGTTGTGGGACGTG  
 1741 CysProThrAspCysPheArgLysHisProAspAlaThrTyrSerArgCysGlySerGly  
 TGCCCCACTGATTGCTTCCGCAAGCATCCGGACGCCACATACTCTCGGTGCGGCTCCGGT  
 ACGGGGTGACTAACGAAGCGTTCGTAGGCCTGCGGTGTATGAGAGCCACGCCGAGGCCA  
 1801 ProTrpIleThrProArgCysLeuValAspTyrProTyrArgLeuTrpHisTyrProCys  
 CCCTGGATCACACCCAGGTGCCTGGTGCAGTACCCGTATAGGCTTTGGCATTATCCTTGT  
 GGGACCTAGTGTGGGTCCACGGACCAGCTGATGGGCATATCCGAAACCGTAATAGGAACA  
 1861 ThrIleAsnTyrThrIlePheLysIleArgMetTyrValGlyGlyValGluHisArgLeu  
 ACCATCAACTACACCATATTTAAATCAGGATGTACGTGGGAGGGGTGGAACACAGGCTG  
 TGGTAGTTGATGTGGTATAAATTTAGTCCTACATGCACCCTCCCAGCTTGTGTCCGAC  
 1921 GluAlaAlaCysAsnTrpThrArgGlyGluArgCysAspLeuGluAspArgAspArgSer  
 GAAGCTGCCTGCAACTGGACGCGGGGCGAACGTTGCGATCTGGAAGACAGGGACAGGTCC  
 CTTGACGGACGTTGACCTGCGCCCCGCTTGCAACGCTAGACCTTCTGTCCCTGTCCAGG  
 1981 GluLeuSerProLeuLeuLeuThrThrThrGlnTrpGlnValLeuProCysSerPheThr  
 GAGCTCAGCCCGTTACTGCTGACCACTACACAGTGGCAGGTCTCCCGTGTCTCTTACA  
 CTCGAGTCGGGCAATGACGACTGGTGTGTGTACCGTCCAGGAGGGCACAAGGAAGTGT  
 2041 ThrLeuProAlaLeuSerThrGlyLeuIleHisLeuHisGlnAsnIleValAspValGln  
 ACCCTACCAGCCTTGTCCACCGGCCTCATCCACCTCCACCAGAACATTGTGGACGTGCAG  
 TGGGATGGTCGGAACAGGTGGCCGGAGTAGGTGGAGGTGGTCTTGTAACACCTGCACGTC  
 2101 TyrLeuTyrGlyValGlySerSerIleAlaSerTrpAlaIleLysTrpGluTyrValVal  
 TACTTGTTACGGGTGGGGTCAAGCATCGCTCTGGGCCATTAAAGTGGGAGTACGTCGTT  
 ATGAACATGCCCCACCCAGTTCGTAGCGCAGGACCCGGTAATTCACCTCATGCAGCAA  
 2161 LeuLeuPheLeuLeuLeuAlaAspAlaArgValCysSerCysLeuTrpMetMetLeuLeu  
 CTCTGTTCTTCTGCTTGCAGACGCGCGCTCTGCTCTGCTTGTGGATGATGCTACTC  
 GAGGACAAGGAAGACGAACGTCTGCGCGCGCAGACGAGGACGAACACCTACTACGATGAG

FIGURE 12-3

2221 IleSerGlnAlaGluAlaAlaLeuGluAsnLeuValIleLeuAsnAlaAlaSerLeuAla  
 ATATCCCAAGCGGAGGCGGCTTTGGAGAACCCTCGTAATACTTAATGCAGCATCCCTGGCC  
 TATAGGGTTCCGCTCCGCCGAAACCTCTTGAGCATTATGAATTACGTCGTAGGGACCGG  
 2281 GlyThrHisGlyLeuValSerPheLeuValPhePheCysPheAlaTrpTyrLeuLysGly  
 GGGACGCACGGTCTTGATCCTTCCTCGTGTCTTCTGCTTTCATGGTATTGTAAGGGT  
 CCTGCGTGCCAGAACATAGGAAGGAGCACAGAAGACGAAACGTACCATAAACTTCCCA  
 2341 LysTrpValProGlyAlaValTyrThrPheTyrGlyMetTrpProLeuLeuLeuLeuLeu  
 AAGTGGGTGCCCCGAGCGGTCTACACCTTCTACGGGATGTGGCCTCTCTCTGCTCCTG  
 TTCACCCACGGGCTCGCCAGATGTGAAGATGCCCTACACCGGAGAGGAGGACGAGGAC  
 2401 LeuAlaLeuProGlnArgAlaTyrAlaLeuAspThrGluValAlaAlaSerCysGlyGly  
 TTGGCGTTGCCCCAGCGGGCGTACGCGCTGGACACGGAGGTGGCCGCGTCTGTGTGGCGGT  
 AACCGCAACGGGGTCCGCCGATGCGCGACCTGTGCCTCCACCGGCGCAGCACACCGCCA  
 2461 ValValLeuValGlyLeuMetAlaLeuThrLeuSerProTyrTyrLysArgTyrIleSer  
 GTTGTTCTCGTGGGTGATGGCGCTGACTCTGTCCACATATTACAAGCGCTATATCAGC  
 CAACAAGAGCAGCCCACTACCGGACTGAGACAGTGGTATAATGTTCCGATATAGTCG  
 2521 TrpCysLeuTrpTrpLeuGlnTyrPheLeuThrArgValGluAlaGlnLeuHisValTrp  
 TGGTGGTGTGGTGGCTTCAGTATTTCTGACCAGAGTGAAGCGCAACTGCACGTGTGG  
 ACCACGAACACCAGGAGTCATAAAGACTGGTCTCACCTTCGCGTTGACGTGCACACC  
 2581 IleProProLeuAsnValArgGlyGlyArgAspAlaValIleLeuLeuMetCysAlaVal  
 ATCCCCCCTCAACGTCCGAGGGGGCGCGACGCCGTCTTACTCATGTGTGTGTA  
 TAAGGGGGGAGTTGCAGGCTCCCCCGCGCTGCGGCAGTAGAATGAGTACACACGACAT  
 2641 HisProThrLeuValPheAspIleThrLysLeuLeuLeuAlaValPheGlyProLeuTrp  
 CACCCGACTCTGGTATTGACATCACCAAAATGCTGCTGGCCGTCTTCGGACCCCTTTGG  
 GTGGGCTGAGACCATAAACTGTAGTGGTTTAACGACGACCGGCAGAAGCCTGGGGAAACC  
 2701 IleLeuGlnAlaSerLeuLeuLysValProTyrPheValArgValGlnGlyLeuLeuArg  
 ATTCTTCAAGCCAGTTTGCTTAAAGTACCCTACTTTGTGCGCGTCCAAGGCCCTTCTCCGG  
 TAAGAAGTTCGGTCAAACGAATTCATGGGATGAAACACGCGCAGGTTCCGGAAGAGGCC  
 2761 PheCysAlaLeuAlaArgLysMetIleGlyGlyHisTyrValGlnMetValIleIleLys  
 TTCTGCGCGTTAGCGCGGAAGATGATCGGAGGCCATTACGTGCAAATGGTCATCATTAAG  
 AAGACGCGCAATCGCGCTTCTACTAGCCTCCGGTAATGCACGTTTACCAGTAGTAATTC  
 2821 LeuGlyAlaLeuThrGlyThrTyrValTyrAsnHisLeuThrProLeuArgAspTrpAla  
 TTAGGGGCGCTTACTGGCACCTATGTTTATAACCATCTCACTCCTCTTCGGGACTGGGCG  
 AATCCCCGCAATGACCGTGGATACAAATATTGGTAGAGTGAGGAGAAGCCCTGACCCGC  
 2881 HisAsnGlyLeuArgAspLeuAlaValAlaValGluProValValPheSerGlnMetGlu  
 CACAACGGCTTGCGAGATCTGGCCGTGGCTGTAGAGCCAGTCGTCTTCTCCCAAATGGAG  
 GTGTTGCCGAACGCTCTAGACCGGCACCGACATCTCGGTGAGCAGAAGAGGGTTTACCTC  
 2941 ThrLysLeuIleThrTrpGlyAlaAspThrAlaAlaCysGlyAspIleIleAsnGlyLeu  
 ACCAAGCTCATCACGTGGGGGGCAGATACCGCCGCGTGGGTGACATCATCAACGGCTTG  
 TGGTTCGAGTAGTGACCCCCCGTCTATGGCGGCGCACGCCACTGTAGTAGTTGCCGAAC  
 3001 ProValSerAlaArgArgGlyArgGluIleLeuLeuGlyProAlaAspGlyMetValSer  
 CCTGTTTCCGCCCCGAGGGGGCAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCC  
 GGACAAAGGCGGGCGTCCCGGCCCTCTATGACGAGCCCGGTGCGCTACCTTACCAGAGG

FIGURE 12-4

3061 LysGlyTrpArgLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeu  
 AAGGGGTGGAGGTTGCTGGCGCCCATCACGGCGTACGCCCAGCAGACAAAGGGGCCTCCTA  
 TTCCCCACCTCCAACGACCGGGGTAGTGCCGCATGCGGGTCTGTTCCCCGGAGGAT

3121 GlyCysIleIleThrSerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGln  
 GGGTGCATAATCACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGTGAGGTCCAG  
 CCCACGTATTAGTGGTCGGATTGACCGGCCCTGTTTTTGGTTCACCTCCCACTCCAGGTC

3181 IleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThr  
 ATTGTGTCAACTGCTGCCAAACCTTCTGGCAACGTGCATCAATGGGGTGTGCTGGACT  
 TAACACAGTTGACGACGGGTTTGAAGGACCGTTGCACGTAGTTACCCACACGACCTGA

3241 ValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMet  
 GTCTACCACGGGGCCGGAACGAGGACCATCGCGTCACCCAAGGGTCTGTGCATCCAGATG  
 CAGATGGTGCCCCGGCCTTGCTCCTGGTAGCGCAGTGGGTTCACGAGACAGTAGGTCTAC

3301 TyrThrAsnValAspGlnAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeu  
 TATACCAATGTAGACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTG  
 ATATGGTTACATCTGGTTCTGGAACACCCGACCGGGCGAGGCGTTCATCGGCGAGTAAC

3361 ThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIle  
 ACACCCTGCACTTGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATT  
 TGTGGGACGTGAACGCCGAGGAGCCTGGAATGGACCAGTGCTCCGTGCGGCTACAGTAA

3421 ProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyr  
 CCCGTGCGCCCGGGGGTGATAGCAGGGGCAGCCTGCTGTCGCCCCGGGCCATTTCTAC  
 GGGCAGCGGGCCGCCACTATCGTCCCCGTCGGACGACAGCGGGCGGGGTAAAGGATG

3481 LeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePhe  
 TTGAAAGGCTCTCGGGGGGTCCGCTGTGTGCCCCGCGGGGCACGCCGTGGGCATATTT  
 AACTTTCCGAGGAGCCCCCAGGCGACAACACGGGGCGCCCCGTGCGGCACCCGTATAAA

3541 ArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheIleProValGluAsn  
 AGGGCCGCGGTGTGCACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAAC  
 TCCCGGCGCCACACGTGGGCACCTACCGATTCCGCCACCTGAAATAGGGACACCTCTTG

3601 LeuGluThrThrMetArgSerProValPheThrAspAsnSerSerProProValValPro  
 CTAGAGACAACCATGAGGTCCCCGGTGTTCACGGATAACTCCTCTCCACCAAGTAGTGCCC  
 GATCTCTGTTGGTACTCCAGGGGCCACAAGTGCTATTGAGGAGAGGTGGTCATCACGGG

3661 GlnSerPheGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysVal  
 CAGAGCTTCCAGGTGGCTCACCTCCATGCTCCACAGGCAGCGGCAAAAGCACCAAGGTC  
 GTCTCGAAGGTCCACCGAGTGGAGGTACGAGGGTGTCCGTGCGCGTTTTCTGTGTTCCAG

3721 ProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAla  
 CCGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGCTGCA  
 GGCCGACGTATACGTGAGTCCCGATATTCCACGATCATGAGTTGGGGAGACAACGACGT

3781 ThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThr  
 ACACTGGGCTTTGGTGTCTTACATGTCCAAGGCTCATGGATCGATCCTAACATCAGGACC  
 TGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTAGGATTGTAGTCTCTGG

3841 GlyValArgThrIleThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeu  
 GGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACCTACGGCAAGTTCCTT  
 CCCCACCTCTTGTTAATGGTGACCGTCGGGGTAGTGATGAGGTGGATGCCGTTCAAGGAA

FIGURE 12-5

7901 AlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSer  
 GCCGACGGCGGGTGTCTCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCC  
 CGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAACTGCTCACGGTGAGG  
 3961 ThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGly  
 ACGGATGCCACATCCATCTTGGGCATCGGCACTGTCTTGACCAAGCAGAGACTGCGGGG  
 TGCTACGGTGTAGGTAGAACCCGTAGCCGTGACAGGAAGTGGTTCGTCTGTACGCCCC  
 4021 AlaArgLeuValValLeuAlaThrAlaThrProProGlySerValThrValProHisPro  
 GCGAGACTGGTGTGTCTCGCCACCGCCACCCTCCGGGCTCCGTCACTGTGCCCCATCCC  
 CGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCGAGGCAGTGACACGGGGTAGGG  
 4081 AsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIle  
 AACATCGAGGAGGTGTCTGTCTCACCACCGGAGAGATCCCTTTTACGGCAAGGCTATC  
 TTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGAAAAATGCCGTTCCGATAG  
 4141 ProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCys  
 CCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTCTGTCAATTCAAAGAAGAAGTGC  
 GGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTAAGTTTCTTCTTCAGG  
 201 AspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaValAlaTyrTyrArgGly  
 GACGAACTCGCCGCAAGCTGGTTCGCATTGGGCATCAATGCCGTGGCCTACTACCGCGGT  
 CTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCACCGGATGATGGCGCCA  
 4261 LeuAspValSerValIleProThrSerGlyAspValValValAlaThrAspAlaLeu  
 CTTGACGTGTCCGTATCCCGACCGCGCATGTTGTCTGTGGCAACCGATGCCCTC  
 GAACTGCACAGGCAGTAGGGCTGGTCCCGCTACAACAGCAGCACCGTTGGCTACGGGAG  
 4321 MetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnThrCysValThrGln  
 ATGACCGGCTATACCGCGCACTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCCAG  
 TACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTATGCACACAGTGGGTC  
 4381 ThrValAspPheSerLeuAspProThrPheThrIleGluThrIleThrLeuProGlnAsp  
 ACAGTCGATTTTCAGCCTTGACCCTACCTCACCATTGAGACAATCACGCTCCCCCAGGAT  
 TGTCACTAAAGTCGGAAGTGGGATGGAAGTGGTAAGTCTGTTAGTGCGAGGGGGTCTTA  
 4441 AlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArg  
 GCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGAAGCCAGGCATCTACAGA  
 CGACAGAGGGCGTGAGTTGCAGCCCCGCTCTGACCGTCCCCCTTCGGTCCGTAGATGTCT  
 4501 PheValAlaProGlyGluArgProSerGlyMetPheAspSerSerValLeuCysGluCys  
 TTTGTGGCACCAGGGGAGCGCCCTCCGGCATGTTGCACTCGTCCGTCTCTGTGAGTGC  
 AAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGGCAGGAGACACTCAGC  
 4561 TyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArg  
 TATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGA  
 ATACTGCGTCCGACACGAACCATACTCGAGTGGGGCGGCTCTGATGTCAATCCGATGCT  
 4621 AlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGly  
 GCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGGAGGGC  
 CGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTTGGTAGAACTTAAACCTCCCG  
 4681 ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGly  
 GTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCAGACAAGCAGAGTGGG  
 CAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTCTGTTTTCGTCTCACC

FIGURE 12-6

4741 GluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaPro  
 GAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCCT  
 CTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCATCCCGAGTTCGGGGA  
 ProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGly  
 4801 CCCCCATCGTGGGACCAGATGTGGAAGTGTTCGCTCAAGCCCACCTCCATGGG  
 GGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTCGGGTGGGAGGTACCC  
 ProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIleThrLeuThrHisPro  
 4861 CCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAAATCACCTGACGCACCCA  
 GGTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAGTGGGACTGCGTGGGT  
 ValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluValValThrSerThrTrp  
 4921 GTCACCAAATACATCATGACATGCATGTGCGCCGACCTGGAGGTGCTCAGGAGCACCTGG  
 CAGTGGTTTATGTAGTACTGTACGTACGACCGGCTGGACCTCCAGCAGTGTCTGTGGACC  
 ValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysVal  
 4981 GTGCTCGTTGGCGGCGTCTCGGCTGCTTTGGCCGCGTATTGCTGTCAACAGGCTGCGTG  
 CAGGAGCAACCGCCGAGGACCGACGAAACCGCGCATAACGGACAGTTGTCCGACGCAC  
 ValIleValGlyArgValValLeuSerGlyLysProAlaIleIleProAspArgGluVal  
 5041 GTCATAGTGGGCAGGGTCGTCTGTTCGGGAAGCCGCAATCATACCTGACAGGGAAGTC  
 CAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTATGGACTGTCCCTTCAG  
 LeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeuProTyrIleGluGln  
 5101 CTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTACCGTACATCGAGCAA  
 GAGATGGCTCTCAAGCTACTCTACCTTCTCAGAGAGTCTGAATGGCATGTAGCTCGTT  
 GlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeuLeuGlnThrAlaSer  
 5161 GGGATGATGCTCGCCGAGCAGTTCAGCAGAAAGGCCCTCGGCCTCCTGCAGACCGCGTCC  
 CCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGAGGAGCTGTGGCGCAGG  
 ArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGlnLysLeuGluThrPhe  
 5221 CGTCAGGCAGAGGTATCGCCCCGTGTCTCAGACCAACTGGCAAAACTCGAGACCTTC  
 GCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTTTTTGAGCTCTGGAAG  
 TrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeuAlaGlyLeuSerThr  
 5281 TGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATAACAATACTTGGCGGGCTGTGCAACG  
 ACCCGCTTCGTATACACCTTGAAGTAGTACCCCTATGTTATGAACCGCCCGAACAGTTGC  
 LeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAlaAlaValThrSerPro  
 5341 CTGCCTGGTAACCCGCCATTGCTTCATTGATGGCTTTTACAGCTGCTGTACCAGCCCA  
 GACGGACCATTGGGGCGGTAACGAAGTAACCTACCGAAAATGTCGACGACAGTGGTGGGT  
 LeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrpValAlaAlaGlnLeu  
 5401 CTAACCACTAGCCAAACCCTCTCTTCAACATATTGGGGGGGTGGGTGGCTGCCAGCTC  
 GATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACCACCGACGGGTGAG  
 AlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAlaGlyAlaAlaIleGly  
 5461 GCCCCCCCGGTGCGGCTACTGCTTTGTGGGCGCTGGCTTAGCTGGCGCGCGCATCGGC  
 CGGCGGGGGCCACGGCGATGACGGAACACCCGCGACCGAATCGACCGGGCGGTAGCCG  
 SerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyrGlyAlaGlyValAla  
 5521 AGTGTGGACTGGGGAAGGTCTCATAGACATCCTTGACGGGTATGGCGCGGGCGTGGCG  
 TCACAACCTGACCCCTTCAGGAGTATCTGTAGGAACGTCCCATACCGCGCCCGCACCGC

FIGURE 12-7

5581 GlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSerThrGluAspLeuVal  
 GGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCCACGGAGGACCTGGTC  
 CCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGGTGCCTCCTGGACCAG  
 5641 AsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGlyValValCysAlaAla  
 AATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGCTGGTCTGTGCAGCA  
 TTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCGACCAGACACGTCGT  
 5701 IleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrpMetAsnArgLeuIle  
 ATACTGCGCCGGCACGTGGCCCGGGCAGGGGGCAGTGCAGTGGATGAACCGGCTGATA  
 TATGACGCGGCGGTGCAACCGGGCCGCTCCCCCGTCACGTCACTTGGCCGACTAT  
 5761 AlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrValProGluSerAspAla  
 GCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTGCCGGAGAGCGATGCA  
 CGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGTGCGTGATGCAGGCCCTCTCGCTACGT  
 5821 AlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGlnLeuLeuArgArgLeu  
 GCTGCCCCGCGTCAC TGCCATACTCAGCAGCCTCACTGTAACCCAGCTCCTGAGGCGACTG  
 CGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTGCGAGGACTCCGCTGAC  
 5881 HisGlnTrpIleSerSerGluCysThrThrProCysSerGlySerTrpLeuArgAspIle  
 CACCAGTGGATAAGCTCGGAGTGTAACCACTCCATGCTCCGGTTCCTGGCTAAGGGACATC  
 GTGGTCACCTATTCGAGCCTCATGTGGTGAGGTACGAGGCCAAGGACCGATTCCCTGTAG  
 5941 TrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeuLysAlaLysLeuMet  
 TGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTAAAAGCTAAGCTCATG  
 ACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGATTTTCGATTGAGTAC  
 6001 ProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyrLysGlyValTrpArg  
 CCACAGCTGCCTGGATCCCCTTTGTGTCTGCCAGCGCGGGTATAAGGGGGTCTGGCGA  
 GGTGTGACGCGACCCCTAGGGGAACACAGGACGGTTCGCGCCCATATTCCCCCAGACCGCT  
 6061 ValAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIleThrGlyHisValLys  
 GTGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAA  
 CACCTGCGGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAGTGACCTGTACAGTTT  
 6121 AsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMetTrpSerGlyThrPhe  
 AACGGGACGATGAGGATCGTCGGTCTTAGGACCTGCAGGAACATGTGGAGTGGGACCTTC  
 TTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGACACCTCACCTGGAAG  
 6181 ProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAlaProAsnTyrThrPhe  
 CCCATTAAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCTGCGCCGAACTACACGTTT  
 GGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGCGGCTTGATGTGCAAG  
 6241 AlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGlnValGlyAspPheHis  
 GCGCTATGGAGGGTGTCTGCAGAGGAATATGTGGAGATAAGGCAGGTGGGGGACTTCCAC  
 CGCGATACTCCACAGACGTCTCCTTATACACCTCTATTCCGTCCACCCCTGAAGGTG  
 6301 TyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGlnValProSerProGlu  
 TACGTGACGGGTATGACTACTGACAATCTCAAATGCCCGTGCCAGGTCCCATCGCCCGAA  
 ATGCACTGCCCATACTGATGACTGTTAGAGTTTACGGGCACGGTCCAGGGTAGCGGGCTT  
 6361 PhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaProProCysLysProLeu  
 TTTTTCACAGAATTGGACGGGGTGCGCTACATAGGTTTGGCCCCCTGCAAGCCCTTG  
 AAAAAGTGCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGGGGGACGTTTCGGGAAC

FIGURE 12-8



6421 LeuArgGluGluValSerPheArgValGlyLeuHisGluTyrProValGlySerGlnLeu  
 CTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCGGTAGGGTCGCAATTA  
 GACGCCCTCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGCCATCCAGCGTTAAT  
 6481 ProCysGluProGluProAspValAlaValLeuThrSerMetLeuThrAspProSerHis  
 CTTTGGCAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTCACTGATCCCTCCCAT  
 GGAACGCTCGGGCTTGGCCTGCACCGGCACAACGCAGGTACGAGTGACTAGGGAGGGTA  
 6541 IleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerProProSerValAlaSer  
 ATACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATACCCCCCTCTGTGGCCAGC  
 TATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTTAGTGGGGGAGACACCGGTCC  
 6601 SerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCysThrAlaAsnHisAsp  
 TCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTGCACCGCTAACCATGAC  
 AGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTTCCGTTGAACGTGGCGATTGGTACTG  
 6661 SerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGlnGluMetGlyGlyAsn  
 TCCCCTGATGCTGAGCTCATAGAGCAACCTCTATGGAGGCAGGAGATGGCGGGCAAC  
 AGGGGACTACGACTCGAGTATCTCCGTTGGAGGATACCTCCGTCTCTACCCGCCGTTG  
 6721 IleThrArgValGluSerGluAsnLysValValIleLeuAspSerPheAspProLeuVal  
 ATCACCAGGGTTGAGTCAGAAACAAAGTGGTGATTCTGGACTCCTTCGATCCGCTTGTG  
 TAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGGAAGCTAGGCGAACAC  
 6781 AlaGluGluAspGluArgGluIleSerValProAlaGluIleLeuArgLysSerArgArg  
 GCGGAGGAGGACGAGCGGAGATCTCCGTACCCGAGAAATCTGCGGAAGTCTCGGAGA  
 CGCTCTCTCTGCTCGCCCTCTAGAGGCATGGCGCTCTTTAGGACGCCTTCAGAGCCTCT  
 6841 PheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnProProLeuValGluThr  
 TTCGCCCAGGCCCTGCCGTTTGGGCGCGGCGGACTATAACCCCCGCTAGTGAGACG  
 AAGCGGGTCCGGGACGGGCAAACCCGCGCGGCTGATATTGGGGGGCGATCACCTCTGC  
 6901 TrpLysLysProAspTyrGluProProValValHisGlyCysProLeuProProProLys  
 TGGAAAAAGCCCGACTACGAACCACTGTGGTCCATGGCTGTCGCTTCCACCTCCAAAG  
 ACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACAGGCGAAGGTGGAGGTTTC  
 6961 SerProProValProProProArgLysLysArgThrValValLeuThrGluSerThrLeu  
 TCCCCTCTGTGCCTCCGCCTCGAAGAAGCGGACGGTGGTCTCACTGAATCAACCCTA  
 AGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAGTGACTTAGTTGGGAT  
 7021 SerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSerSerThrSerGlyIle  
 TCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCCTCACTTCCGGCATT  
 AGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGGAGTTGAAGGCCGTAA  
 7081 ThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGlyCysProProAspSer  
 ACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGCTGCCCCCGACTCC  
 TGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCGACGGGGGGGCTGAGG  
 7141 AspAlaGluSerTyrSerSerMetProProLeuGluGlyGluProGlyAspProAspLeu  
 GACGCTGAGTCTATTCTCCATGCCCCCTGGAGGGGGAGCCTGGGGATCCGGATCTT  
 CTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGACCCCTAGGCCTAGAA  
 7201 SerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGluAspValValCysCys  
 AGCGACGGGTATGGTCAACGGTCAGTAGTGAGGCCAACCGGAGGATGTCGTGTGTGTC  
 TCGCTGCCAGTACCAGTTGCCAGTCATCTCCGGTTGCGCCTCTACAGCACACGACG

FIGURE 12-9

SerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAlaAlaGluGluGlnLys  
 7261 TCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCCGCGGAAGAACAGAAA  
 AGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGCGCCTTCTGTCTTT  
 LeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsnLeuValTyrSerThr  
 7321 CTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAATTGGTGTATTCCACC  
 GACGGGTAGTTACGTGATTCTGAGCAACGATGCAGTGGTGTAAACCACATAAGGTGG  
 ThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAspArgLeuGlnValLeu  
 7381 ACCTCACGCAGTGCTTGCCAAAGGCAGAAAGTACATTTGACAGACTGCAAGTTCTG  
 TGGAGTGGTTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTGTCTGACGTTCAAGAC  
 AspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAlaSerLysValLysAla  
 7441 GACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGCGCTCAAAGTGAAGGCT  
 CTGTCGGTAATGGTCTGCATGAGTTCTTCCAATTTCTGTCGCCGAGTTTTCACTTCCGA  
 AsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHisSerAlaLysSerLys  
 7501 AACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACTCAGCCAAATCCAAG  
 TTGAACGATAGGCATCTCCTTCGAACGTGCGACTGCGGGGGTGTGAGTCGGTTTAGGTTT  
 PheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAlaValThrHisIleAsn  
 7561 ITTGGTTATGGGGCAAAGACGTCCGTGCGCATGCCAGAAAGGCCGTAAACCCACATCAAC  
 AAACCAATACCCCGTTTCTGCAAGCAACGGTACGGTCTTTCGGCATTTGGGTGTAGTTG  
 SerValTrpLysAspLeuLeuGluAspAsnValThrProIleAspThrThrIleMetAla  
 7621 TCCGTGTGGAAGACCTTCTGGAAGACAATGTAACACCAATAGACACTACCATCATGGCT  
 AGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTGTGATGGTAGTACCGA  
 LysAsnGluValPheCysValGlnProGluLysGlyGlyArgLysProAlaArgLeuIle  
 7681 AAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGTTCGTAAGCCAGCTCGTCTCATC  
 TTCTTGCTCCAAAAGACGCAAGTTCGACTCTTCCCCCAGCATTCGGTCGAGCAGAGTAG  
 ValPheProAspLeuGlyValArgValCysGluLysMetAlaLeuTyrAspValValThr  
 7741 GTGTTCCCCGATCTGGGCGTGCGGTGTGCGAAAAGATGGCTTTGTACGACGTGGTTTACA  
 CACAAGGGGCTAGACCCGACGCGCACAGCTTTTCTACCGAAACATGCTGCACCAATGT  
 LysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyrSerProGlyGlnArg  
 7801 AAGCTCCCCCTGGCCGTGATGGGAAGCTCCTACGGATTCCAATACTCACCAGGACAGCGG  
 TTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAGGTTATGAGTGGTCTGTCCGC  
 ValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMetGlyPheSerTyrAsp  
 7861 GTTGAATTCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATGGGGTTCTCGTATGAT  
 CAACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTGGGGTTACCCCAAGAGCATACTA  
 ThrArgCysPheAspSerThrValThrGluSerAspIleArgThrGluGluAlaIleTyr  
 7921 ACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACGGAGGAGGCAATCTAC  
 TGGGCGACGAACTGAGGTGTCAGTACTCTCGCTGTAGGCATGCCTCCTCCGTTAGATG  
 GlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSerLeuThrGluArgLeu  
 7981 CAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCCCTCACCAGAGGCTT  
 GTTACAACACTGGAGCTGGGGGTTCCGGCGCACCGGTAGTTTCAGGGAGTGGCTCTCCGAA  
 TyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGlyTyrArgArgCysArg  
 8041 TATGTTGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGCTATCGCAGGTGCCGC  
 ATACAACCCCGGAGAAATGGTTAAGTTCCCCCTCTTGACGCCGATAGCGTCCACGGCG

FIGURE 12-10

8101 AlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCysTyrIleLysAlaArg  
 GCGAGCGGCGTACTGACAAC TAGCTGTGGTAACACCTCACTTGCTACATCAAGGCCGG  
 CGCTCGCCGCATGACTGTTGATCGACACCATGTGGGAGTGAACGATGTAGTTCGGGGCC  
 8161 AlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuValCysGlyAspAspLeu  
 GCAGCCTGTGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTGTGGCGACGACTTA  
 CGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCACACACCGCTGCTGAAT  
 8221 ValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSerLeuArgAlaPheThr  
 GTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCGCGAGCCTGAGAGCCTTCACG  
 CAGCAATAGACACTTTCGCGCCCCCAGGTCTCTGCGCGCTCGGACTCTCGGAAGTGC  
 8281 GluAlaMetThrArgTyrSerAlaProProGlyAspProProGlnProGluTyrAspLeu  
 GAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCAACAGAGCTGGAAGAGG  
 CTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGTGGTCTTATGCTGAAC  
 8341 GluLeuIleThrSerCysSerSerAsnValSerValAlaHisAspGlyAlaGlyLysArg  
 GAGCTATAACATCATGCTCCTCCAACGTGTCAGTCGCCACGACGCGCTGGAAAGAGG  
 CTCGAGTATTGTAGTACGAGGAGGTGCACAGTCAGCGGGTGTGCCGCGACCTTCTCC  
 8401 ValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAlaAlaTrpGluThrAla  
 GTCTACTACCTCACCGTGACCTTACAACCCCTCGCGAGAGCTGCGTGGGAGACAGCA  
 CAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGACGCACCTCTGTCTG  
 8461 ArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPheAlaProThrLeuTrp  
 AGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTTGGCCCCACAGCTTGG  
 TCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAAACGGGGGTGTGACACC  
 8521 AlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAlaArgAspGlnLeuGlu  
 GCGAGGATGATACTGATGACCCATTCTTTAGCGTCCTTATAGCCAGGACAGCTTGAA  
 CGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGGTCCCTGGTCAACTT  
 8581 GlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGluProLeuAspLeuPro  
 CAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAACCACTTGATCTACCT  
 GTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTTGGTGAAGTAGATGGA  
 8641 ProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHisSerTyrSerProGly  
 CCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCACAGTTACTCTCCAGGT  
 GGTTAGTAAGTTTCTGAGGTACCGAGTCGCGTAAAAGTGAGGTGTCAATGAGAGGTCCA  
 8701 GluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValProProLeuArgAlaTrp  
 GAAATTAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCGCCCTTGCGAGCTTGG  
 CTTTAATTATCCACCGCGTACGGAGTCTTTTGAACCCCATGGCGGGAACGCTCGAACC  
 8761 ArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGlyGlyArgAlaAlaIle  
 AGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGAGGCAGGGCTGCCATA  
 TCTGTGGCCCGGGCCTCGCAGGCGGATCCGAAGACCGGTCTCTCCGTCCCGACGGTAT  
 8821 CysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
 TGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAAC  
 ACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTTG

FIGURE 12-11

EP 0 419 182 B1

11-1519

primer J150S

J1 ACTGCCCTGAACTGCAATGA

PT G

SerLeuLysThrGlyPheLeuAlaAlaLeuPheTyrThrHisLysPheAsnAlaSerG1  
 J1 1 CTCCCTCAAACTGGGTTTCCTGCCCGCTGTTCTACACACAGTTCAACGGCTCCGG 90  
 PT TAG C C C GGT G AG T TCAC T T A  
Asn Trp Gly His Ser  
 primer 188A for J1-1216

yCysProGluArgMetAlaSerCysArgSerIleAspLysPheAspGlnGlyTrpGlyPr  
 J1 61 ATGCCCGGAGCGCATGCCAGCTGTGCTCCATTGACAAGTTCCAGCGGGATGGGGTCC 120  
 PT C T T A GC A C AC C AC G T T C C  
 Leu ProLeuThrAsp

cIleThrTyrAlaGlnProAspAsnSerAspGlnArgProTyrCysTrpHisTyrAlaPr  
 J1 121 CATCACCCTATGCTCAACCTGACAACCTCCACCACAGGCCCTATTGCTGGCACTACGCACC 180  
 PT T GT CA CGGAAG GG C C C C C C  
 Ser AsnGlySerGlyPro Pro

oArgGlnCysGlyIleValProAlaSerGlnValCysGlyProValTyrCysPheThrPr  
 J1 181 TCGACAGTGTGGTATCGTACCCCGCTCGCAGGTGTGCGGTCCAGTGTATTGCTTCACCCC 240  
 PT AAA CT C T G AA AGT T G A T  
 LysPro LysSer

oSerProValValValGlyThrThrAspArgPheGlyAlaProThrTyrAsnTrpGlyAs  
 J1 241 AAGCCCTCTTGTAGTGGGGACGACCGATCGTTTCGGCGCCCCCTACGTATAACTGGGGCCA 300  
 PT C C G G A CA G CG G C C C G T  
 Ser Ser Gl

EAsnGluThrAspValLeuLeuLeuAsnAsnThrArgProProHisGlyAsnTrpPheG1  
 J1 301 CAATGAGACGGACCTGCTGCTCCTAAACAACACGGCGCCCCCGACGGCAACTGGTTCGG 360  
 PT A T CT CG T T CA A TG T  
 u Asp PheVal Leu

yCysThr  
 J1 361 CTGTACA 387  
 PT T CTGGATGAACTCAACTGGATT  
 primer 199A

Nucleotide Match: 259/387 (70.6%)  
 Amino Acid Match (stringent): 93/122 (76.2%)  
 (relaxed): 111/122 (91.0%)

Prototype HCV (PT) sequences different from Japanese HCV (J1) are shown.  
 Relaxed amino acid match: Gly=Ala=Pro=Ser=Thr. Asp=Glu, Asn=Gln.  
 Arg=Asp=Lys=His. Leu=Ile=Val=Met. Phe=Trp=Tyr.  
 Underline, different amino acid in relaxed matching.

FIGURE 13

FIGURE 14-1

core to NS1 vs. HCV-1

ProLeuVal  
TCCGCTCGTC  
A-----

HCV-1

11 GlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp  
GGCGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCACATGGTGTCCGGGTCTGGAGGAC  
-----TC-T-A-----G-----C-----A-----

71 GlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla  
GGCGTGAACATATGCAACAGGGAATTGCCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCT  
-----CC-T-T-----TC-----

131 LeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValArgAsnValSerGlyIle  
CTGCTGTCCTGTTTGACCATCCCAGCTTCCGCTTATGAAGTGCACAACGTGTCCGGGATA  
-----C-T-C-----TG-G-C-----G-C-CC-----TCCA-G-----C-T  
Val Gln SerThr Leu

191 TyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAlaAspValIle  
TACCATGTCAAAACGACTGCTCCAACCTCAAGCATGTGTATGAGGCGGCGACGTGATC  
-----C-----C-T-T-----C-T-----G-T-----C-----T-CC-----  
Pro Ala

251 MetHisAlaProGlyCysValProCysValArgGluAsnAsnSerSerArgCysTrpVal  
ATGCATGCCCCCGGTGCGTGCCCTGCGTTCGGGAGAACAAATCTCTCCCGTTGCTGGGTA  
C-----CA-T-G-----C-T-----T-GG-----CG-----GA-G-----A  
Leu Thr Gly Ala

311 AlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrThrLeuArgArg  
GCGCTCACTCCCACGCTCGCGGCCAGGAATGCCAGCGTCCCACTACGACATTACGACGC  
-----A-G-C-T-----G-G-CA-----G-----AAC-----G-G-----CAGC-T-----T  
Met Val Thr AspGlyLysLeu Ala Gln

371 HisValAspLeuLeuValGlyThrAlaAlaPheCysSerAlaMetTyrValGlyAspLeu  
CAGCTCGACTGCTCGTTGGGACGGCTGCTTTCTGCTCCGCTATGTACGTGGGGATCTC  
-----A-TC-----T-C-----GC-CA-CC-----T-G-CC-C-----C-A  
Ile Ser ThrLeu Leu

431 CysGlySerValPheLeuIleSerGlnLeuPheThrPheSerProArgArgHisGluThr  
TGGGATCTGTTTTCTCATCTCCAGCTGTTACCTTCTCGCTCGCCGGCATGAGACA  
-----G-----C-T-TG-GG-----A-----T-CA-G-C-CTG-G  
ValGly Trp

491 ValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMetAlaTrp  
GTACAGGACTGCAACTGCTCAATCTATCCCGGCCACGTATCAGGCCATCGCATGGCTTGG  
ACG--A-GT-----T-----T-----TA--A-G--T-C-----A--  
Thr Gly IleThr

551 AspMetMetMetAsnTrpSerProThrAlaAlaLeuValValSerGlnLeuLeuArgIle  
GATATGATGATGAACCTGGTGGCCAGCGCCTTAGTGGTGTGCGAGTTACTCCGGATC  
-----C-T-----A-G-G-G-AA-G-T-C-G-----  
Thr MetAla

611 ProGlnAlaValMetAspMetValAlaGlyAlaHisTrpGlyValLeuAlaGlyLeuAla  
CCACAAGCTGTATGACATGGTGGCGGGGGCCACTGGGGAGTCTAGCGGGCCTTGCC  
-----CA-T-----A-C-T-T-----G-----A-A-G  
IleLeu Ile Ile

671 TyrTyrSerMetValGlyAsnTrpAlaLysValLeuIleValMetLeuLeuPheAlaGly  
TACTATTCCATGGTGGGGAACCTGGGCTAAGGTTTGTATGTGATGCTACTCTTGGCGGC

```

--T-TC-----G-----CC-G-A-C-G-A-----
Phe                      Val    Leu

ValAspGlyHisThrArgValThrGlyGlyValGlnGlyHisValThrSerThrLeuThr
J1  731 GTTGACGGGCATACCCGCGTGACGGGGGGGTGCAAGGCCACGTACCTCTACACTCAGC
--C---C-G-A---A---C---AAGTGCC-----ACTGTG---GGAT-TGTT
AlaGlu  His                      SerAla    ThrVal  GlyPheVal

SerLeuPheArgProGlyAlaSerGlnLysIleGlnLeuValAsnThrAsnGlySerTrp
J1  791 TCCCTCTTTAGACCTGGGGCGTCCCGAGAAATTGAGCTTGTAACACCAATGGCAGTTGG
AG---C-CGC---A---C---CAAGC---CG-C---GA-C---C-----
LeuAla                      Lys  AsnVal    Ile

HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuGlnThrGlyPheLeuAlaAla
J1  851 CATATCAACAGGACTGCCCTGAACTGCAATGACTCCCTCCAAACTGGGTTCTTGCCGCG
--CC---T---C---G-----TAG---A-C---C---C-GGT-G---A-G-
Ser                      Asn    Trp    Gly

J1  911 CTGTTCTACACACACAAGTTCAACGCGTCCGGATGCCCGGAGCGCATGGCCAGCTGTCCG
--T---TCACC-----T-T-A-C-T-T---A-GC-A-----C-A
His                      Ser                      Leu

SerIleAspLysPheAspGlnGlyTrpGlyProIleThrTyrAlaGlnProAspAsnSer
J1  971 TCCATTGACAAGTTGACACAGGGATGGGGTCCCATCACCTATGCTCAACCTGACAACTCG
C--C--AC-G-T-T-----C---C---T---GT---CAACGGAAGCGG-C-C
ProLeuThrAsp                      Ser    AsnGlySerGlyPro

AspGlnArgProTyrCysTrpHisTyrAlaProArgGlnCysGlyIleValProAlaSer
J1  1031 GACCAGAGGCCGTATTGCTGGCACTACGCACCTCGACAGTGTGGTATCGTACCCGCGTCG
-----C-C---C---C-----C-C---AAA---CT---C---T---G-----AA-
Pro    LysPro                      Lys

GlnValCysGlyProValTyrCysPheThrProSerProValValValGlyThrThrAsp
J1  1091 CAGGTGTGCGGTCCAGTGTATTGCTTACCCCAAGCCCTGTGTAGTGGGGACGACCGAT
AGT-----T---G-A-----T-C---C---G-G---A-----
Ser

ArgPheGlyAlaProThrTyrAsnTrpGlyAspAsnGluThrAspValLeuLeuLeuAsn
J1  1151 CGTTTCGGCGCCCTACGTATAACTGGGGGACAATGAGACGGACGTGCTGCTCCTAAAC
A-G-CG---G-C---C---G---T-A---T---CT-CG---T---
Ser                      Ser    Glu  Asp    PheVal

AsnThrArgProProHisGlyAsnTrpPheGlyCysThr
J1  1211 AACACGCGGCCCGGCACGGCAACTGGTTCGGCTGTACA
--T---CA---A---TG---T-----T-----
Leu

```

FIGURE 14-2

GlyAsnTrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysThrCysGlyGly  
 TGGGCAACTGGTTCGGCTGTACATGGATGAATAGCACTGGTTCACCAAGACGTGCGGAG  
 -----T-----T-----C-----CTCA-----A-----AGT-----  
 Val  
 ProProCysAsnIleGlyGlyValGlyAsnAsnThrLeuThrCysProThrAspCysPhe  
 61 GCGGGCGTGTAAACATCGGGGGGGTCGGCAACAACACCTGACCTGCCCCACGGACTGCT  
 CG--T--T--GT-----A--CG-----C--CA-----T--T--  
 Val Ala His  
 ArgLysThrProThrAlaThrTyrThrLysCysGlySerGlyProTrpLeuThrProArg  
 121 TCCGGAAGACCCCGACGGCCACTTACACAAATGTGGTTCGGGCCCTTGGTTGACACCTA  
 ---C---CAT---GAC---A---T-TCGG--C--C--C--T--C---A-C-----C-  
 His Asp SerArg Ile  
 CysLeuValAspTyrProTyrArgLeuTrpHisTyrProCysThrValAsnPheThrIle  
 181 GGTGCTTGGTTGACTACCCATACAGGCTCTGGCACTACCCCTGCACTGTCAACTTTACCA  
 ---C---C-----G--T-----T-----T--T--T--T--CA---AC---  
 Ile Tyr  
 PheLysValArgMetTyrValGlyGlyValGluHis  
 241 TCTTCAAGGTTAGGATGTATGTGGGGGCGTGGAGCAC  
 -A--T--AA-C-----C-----A--G--C--A--  
 Ile

FIGURE 15

## C200 region sequence vs. HCV-1

C200  
 HCV-1 3781 AsnMetSerLysAlaHisGlyThrAspProAsnIleArgThr  
 3799 AATATGTCCAAGGCACATGGCACCGACCCCAACATCAGAACT  
 ACACTGGGCTTTGGTGCTT-C-T-G-T-T-T-G-C  
 ThrLeuGlyPheGlyAlaTyr Ile

C200  
 HCV-1 3841 GlyValArgThrIleThrThrGlyAlaProIleThrTyrSerThrTyrArgLysPheLeu  
 GGGGTAAGGACCATCACACAGGTGCCCCATTACGTACTCCACCTATCCCAAGTTCCTT  
 -G-A-A-T-T-CAG-C-CG  
 Ser Gly

C200  
 HCV-1 3901 AlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle  
 GCCGACGGIGGTGTCTCCGGGGCGCCTATGACATCATAA 3940  
 -C-G-G-T-A-TTTGTGACGAGTGCCACTCC  
 IleCysAspGluCysHisSer

HCV-1 3961 ACCGATGCCACATCCATCTTGGGCATCGGCACTGTCTTGACCAAGCAGAGACTGCGGGG  
 ThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGly

HCV-1 4021 GCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTCACGTGCCCCATCCC  
 AlaArgLeuValValLeuAlaThrAlaThrProProGlySerValThrValProHisPro

C200  
 HCV-1 4081 SerIle  
 AACATCGAGGAGGTGTCTGTCCACCACCGGAGAGATCCCTTTTACGGCAAGGCT- 4134 AAGCAIC  
 AsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAla

C200  
 HCV-1 4141 ProIleGluAlaIleLysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCys  
 CCCATCGAGGCCATCAAGGGGGAAGGCATCTCATCTTCTGCCATTCCAAGAAGAAGTGT  
 -C-A-TA-G-A-T-A-C  
 Val

C200  
 HCV-1 4201 AspGluLeuAlaAlaLysLeuSerAlaLeuGlyLeuAsnAlaValAlaTyrTyrArgGly  
 GACGAGCTCGCCGCAAGCTGTACGCCCTCGGACTCAATGCCGTGGCGTATTACCGCGGT  
 -A-GTC-AT-G-CA-C-C  
 Val Ile

FIGURE 16-1



C200 4261 LeuAspValSerValIleProThrSerGlyAspValValValAlaThrAsp  
 HCV-1 CTTGATGTGTCCGTCATACCACTAGCGGAGACGTCGTTGTCGTGGCAACAGACGC 4316  
 -----C-----C-G-C-----C-T-T-C-----C-T-CCTC

HCV-1 4321 ATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAG  
 MetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsnThrCysValThrGln

HCV-1 4381 ACAGTCGATTTACAGCCTTGACCCTACCTTCACCATGAGACAATCAGGCTCCCCCAGGAT  
 ThrValAspPheSerLeuAspProThrPheThrIleGluThrIleThrLeuProGlnAsp

HCV-1 4441 GCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGA  
 AlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArg

HCV-1 4501 TTTGTGGCACCAGGGGAGCGCCCTCCGGCATGTTGACTCGTCCGTCCCTGTGAGTGC  
 PheValAlaProGlyGluArgProSerGlyMetPheAspSerSerValLeuCysGluCys

HCV-1 4561 TATGACGCAGGCTGTGCTTGGTATGAGCTCAGCCCCCGGAGACTACAGTTAGGCTACGA  
 TyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArg

HCV-1 4621 GCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGGAGGGC  
 AlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGly

HCV-1 4681 GTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGG  
 ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGly

HCV-1 4741 GAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCSTGTGCGCTAGGGCTCAAGCCCTT  
 GluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaPro

HCV-1 4801 CCCCCATCGTGGGACCAGATGTGGAAGTGTGATTGCGCTCAAGCCCACCTCCATGGG  
 ProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGly

HCV-1 4861 CCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTAAATCACCCTGACGCACCCA  
 ProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIleThrLeuThrHisPro

FIGURE 16-2

HCV-1 4921 GTCACCAAAIACATCATGACATGCATGTCGGCCGACCTGGAGGTCGTCACGAGCACCTGG  
ValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluValValThrSerThrTrp

HCV-1 4981 GTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCTGTCAACAGGCTGCGTG  
ValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysVal

HCV-1 5041 GTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATACCTGACAGG-----  
ValIleValGlyArgValValLeuSerGlyLysProAlaIleIleProAspArg  
GluVal  
5094 GAAGTC

C200 5101 LeuTyrArgGluPheAspGluMetGluGluCysAlaSerHisLeuProTyrIleGluGln  
HCV-1 CTCTACCGAGAGTTTCGATGAGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAG  
-----T-TCAG-----T-A-G-----G--A  
SerGln

C200 5161 GlyMetGlnLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeuLeuGlnThrAlaThr  
HCV-1 GGAATGCAGCTCGCCGASCAATTCAGCAGAAGGCGCTCGGGTTGCTGC AACAGCCACC  
--G--AT-----G-----C-----CC-C-----G--C--GT--  
Met Ser

C200 5221 LysGlnAlaGluAlaAlaAlaProCysGluSerMetHisAlaSer  
HCV-1 AAGCAAGCGGAGGCTGCTGCTCCGTGTGAGTCAATGCACGCCTCGA  
CGT--G--A-----T-ATC--C--TGC--TCCAG--CCA--TGGCAA-AACTCGAGACCTTC  
Arg ValIle AlaValGlnThrAsnTrpGlnLysLeuGluThrPhe

HCV-1 5281 TGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATA  
TrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln

FIGURE 16-3

**FIGURE 17**

## Core sequence vs. HCV-1

J1 V-1	1	CGGTCTAGCCATGGCGTTAGTATGAGTGTCTGTCAGCCTCCAGGACCCCCCTCCC
J1 HCV-1	57	GGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGGACGACCGGG
J1 HCV-1	117	TCCTTTCTTGGATCAACCCGCTCAATGCCTGGAGATTGGGGCGTGCCCCGCGAGACTGC -----A-----
J1 HCV-1	177	TAGCCGAGTAGTGTGGGTGCGGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTTGCGA
J1 HCV-1	237	GTGCCCCGGGAGGTCTCGTAGACCGTGCATCATGAGCACAAATCCTAAACCTCAAAGAAA -----C-----G-----A----- Lys
J1 V-1	297	ThrLysArgAsnThrAsnArgArgProGlnAspValLysPheProGlyGlyGlyGlnIle AACCAAACGTAACACCAACCGCGCCACAGGACGTCAAGTTCCCGGGCGGTGGTCAGAT --A-----T-----T-----C----- Asn
J1 HCV-1	357	ValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeuGlyValArgAlaThrArg CGTTGGTGGAGTTTACCTGTTGCCGCGCAGGGGCCCCAGGTTGGGTGTGCGCGCGACTAG -----T-----T-----A-----G-----
J1 HCV-1	417	LysThrSerGluArgSerGlnProArgGlyArgArgGlnProIleProLysAlaArgGln GAAGACTTCCGAGCGGTCGCAACCTCGTGGAAGGCGACAACCTATCCCAAGGCTCGCCA A-----A-----T-----A-----T-----G-----T-----G----- Arg
J1 HCV-1	477	ProGluGlyArgAlaTrpAlaGlnProGlyTyrProTrpProLeuTyrGlyAsnGluGly GCCCCAGGGCAGGGCCTGGGCTCAGCCCCGGGTACCCTTGCCCCCTCTATGGCAACGAGGG -----A-----T----- Thr
J1 V-1	537	MetGlyTrpAlaGlyTrpLeu CATGGGGTGGGCAGGATGGCTCCT -TGC-----G----- Cys

FIGURE 18